

## 4 Methode

### 4.1 Untersuchungsgegenstand

Die vorliegende Studie untersucht die Veränderung der interpersonal-affektiven Schemata von Patienten, die sich in psychoanalytische Langzeitbehandlungen begeben haben. Damit konzentrieren wir uns auf das oben aufgeführte Ziel psychoanalytischer Behandlungen, den Patienten befriedigendere Beziehungen zu Anderen zu ermöglichen, und lassen die weiteren erwähnten Zielsetzungen der Psychoanalyse unberücksichtigt. In den Mittelpunkt stellen wir dabei die *subjektiven Beziehungsrepräsentationen* der Patienten.

Die meisten theoretischen Formulierungen des Konzeptes der interpersonellen Schemata beinhalten die Annahme, dass zwar die Schemata selbst größtenteils unbewusst sind, derjenige, der ein Schema „in sich trägt“, dessen *Auswirkungen* aber durchaus hautnah erleben kann: So ist es beispielsweise vorstellbar, dass ein Patient das folgende unbewusste Schema verinnerlicht hat: „Ich bin nicht liebenswert. Es lohnt sich nicht auf andere zuzugehen, weil diese mich ohnehin ablehnen werden.“ Während er zu dieser Grundüberzeugung selbst keinen Zugang mehr hat, kann er jedoch klar benennen, dass es ihm – auf der Ebene des sichtbaren Verhaltens – nicht gelingt, befriedigende Beziehungen zu Anderen aufzubauen, und dass er darunter leidet.

Schemata selbst sind theoretische Konstrukte, lassen sich also niemals beobachten, sondern nur hypothetisch erschließen. Unsere Operationalisierung folgt der Logik, dass die *Nennungshäufigkeiten* bestimmter Beziehungsaspekte in den Erzählungen eines Patienten die Bedeutsamkeit dieser Themata für denselben anzeigen. Im Falle des oben genannten Patienten fiel zum Beispiel auf, wie oft er davon berichtet, von Anderen zurückgewiesen oder missachtet worden zu sein. Wir würden dann davon ausgehen, dass die Häufung derartiger Erfahrungen kein Zufall ist, sondern Ausdruck eines dahinter liegenden Schemas. Das Schema könnte in zweifacher Weise die auffällige Häufung des Ablehnungsthemas bewirken: Erstens, indem der Patient die Verhaltensweisen seiner Objekte unter dem Eindruck des Schemas in verzerrter Art und Weise wahrnimmt, sich zum Beispiel vor allem auf diejenigen Informationen konzentriert, die seine Erwartungen bestätigen, und widersprechende zu wenig berücksichtigt. Zweitens könnte er seine Beziehungen tatsächlich in Entsprechung zum Schema konstellieren, also etwa durch eigenes schroffes, distanziertes oder ängstliches Verhalten tatsächlich Ablehnung auf Seiten der Objekte provozieren.

Auch eine besonders *niedrige* Nennungshäufigkeit eines Beziehungsthemas kann auf Charakteristika des Schema-Repertoires einer Person verweisen: Der Bericht eines Patienten, der im Laufe eines einstündigen Interviews kein einziges Mal davon spricht, von Anderen angenommen, akzeptiert oder bestätigt worden zu sein lässt ähnliche Schlüsse auf die dahinter liegenden Beziehungsrepräsentationen zu wie im oben genannten Fall. Wir betrachten die Kopplung dessen, was ein Patient über sich und seine Objekte berichtet, mit den dahinter vermuteten interpersonal-affektiven Schemata als sehr eng und gehen daher davon aus, dass letztere sich *am besten* auf der Grundlage von Beziehungsnarrativen ermitteln lassen. Damit befinden wir uns auch im Einklang mit den meisten der bereits dargestellten Verfahren zur Erfassung von Beziehungsschemata. Ein Spezifikum unserer Vorgehensweise besteht im Vergleich zu anderen Verfahren jedoch darin, dass theoretisch fundierte Schlussfolgerungen auf tiefere Bedeutungsebenen des Gesagten möglichst weitgehend ausgeschlossen werden sollen. Die Codierung der Wunsch-Komponente im CCRT beispielsweise stellt in den

meisten Fällen eine solche Schlussfolgerung dar, weil Patienten eher selten spontan berichten, was sie sich vom Gegenüber erhofft hätten. Durch die Beschränkung auf diejenigen Bedeutungen, die der Patient selbst dem Gesagten gibt, wollen wir vor allem zwei Effekte erzielen: Zum einen führt eine solche Vorgehensweise mit einiger Wahrscheinlichkeit zu hinreichend reliablen Daten. Zum anderen erleichtert diese Konzentration auf das Offensichtliche es Untersuchern, die sich verschiedenen theoretischen Traditionen der Psychologie verbunden fühlen, den Daten Validität zuzuerkennen.

In der klinischen Behandlungssituation stehen dem Therapeuten im Wesentlichen zwei Informationsquellen zur Verfügung, auf deren Grundlage er beurteilen kann, ob sich in der Beziehungsgestaltung des Patienten systematische Veränderungen ergeben haben: Zum einen seine Gegenübertragung, also sein eigenes Erleben im unmittelbaren Kontakt mit dem Patienten. Zum anderen das, was der Patient über seine Beziehungen zu seinen Objekten berichtet. Beide Bereiche sind gleichermaßen wichtig. In der vorliegenden Arbeit konzentrieren wir uns jedoch auf den letztgenannten, weil dieser eindeutiger dem Patienten zuzuordnen ist, als die Gegenübertragung, bei deren Zustandekommen die Interaktion von Patienten- und Therapeuteneinflüssen eine größere Rolle spielt. Wir verzichten also bewusst auf eine der beiden Informationsquellen, um eine größere Präzision unserer Aussagen zu erzielen, indem diese sich ausschließlich auf das *verbalisierte Erleben des Patienten* beschränken.

An den Beziehungsrepräsentationen der Patienten interessieren uns inhaltlich vor allem die folgenden zwei Bereiche, die wir zunächst für weitgehend unabhängig voneinander halten:

Zum einen ist es wichtig, welche *Handlungen* zwischen den beteiligten Interaktanden stattfinden. Auf dieser Ebene soll all das erfasst werden, was zwischen zwei Menschen passiert und zumindest prinzipiell von Anderen beobachtbar wäre: Eine Person macht zum Beispiel einer anderen einen Vorwurf, worauf die andere Person sich dem weiteren Gespräch entzieht und schmolzt.

Die zweite wichtige Ebene ist die des emotionalen Erlebens: Was *fühlen* die Beteiligten in den beschriebenen sozialen Interaktionen? Im oben genannten Beispiel mag sich die zuletzt genannte Person etwa deshalb zurückziehen, weil sie sich durch den Vorwurf des Anderen verletzt und gekränkt fühlt. Sie mag aber im Gegenteil auch eher so etwas wie Trotz empfinden und deshalb den Kontakt abbrechen. Hier deutet sich bereits die erwähnte prinzipielle Unabhängigkeit der beiden Ebenen an.

## **4.2 Rahmen und Design der Untersuchung**

Die vorliegende Studie wurde im Rahmen der zweiten Förderperiode des *Graduiertenkollegs Klinische Emotionsforschung* von der Deutschen Forschungsgemeinschaft (DFG) gefördert. Das Untersuchungsmaterial entstammt sämtlich dem Datenpool der *Praxisstudie Analytische Langzeittherapie* (PAL; Grande, Rudolf & Oberbracht, 1997) an der Psychosomatischen Universitätsklinik Heidelberg. Die PAL untersucht in einem prospektiven Design die Therapieverläufe von Patientinnen und Patienten, die sich in Heidelberg und Umgebung in psychoanalytische Behandlung begeben haben, wobei als Einschlusskriterium für die Gruppe der Psychoanalysen definiert wurde, dass die Behandlung mindestens dreimal pro Woche auf der Couch liegend durchgeführt wird.

Die Probanden kommen in regelmäßigen Abständen – zunächst viertel- dann halbjährlich – in die Psychosomatische Klinik, wo sie von einem psychoanalytisch geschulten Interviewer zu ihrer aktuellen Lebenssituation interviewt werden. Das Interview wird entweder von einem männlichen oder von einem weiblichen Interviewer durchgeführt, wobei die Zuordnung zufällig erfolgt. Der Interviewer hält seine Eindrücke aus dem Gespräch schriftlich und in standardisierter Form fest, die Videoaufzeichnungen der Gespräche werden von zwei weiteren, unabhängigen Ratern im Hinblick auf persönlichkeitsstrukturelle Veränderungen im Therapieverlauf beurteilt. Im Anschluss an das Gespräch bearbeiten die Patienten einen Satz Fragebögen, vorwiegend zu den Bereichen Symptomatik, Beziehungsschwierigkeiten und Lebenszufriedenheit. Parallel dazu berichten die behandelnden Psychoanalytiker regelmäßig über ihre aktuelle Sicht des analytischen Prozesses, über ihre Behandlungstechnik, sowie über einzelne, ihnen besonders wichtig erscheinende Behandlungsstunden.

Eine zweite Gruppe von Patienten, die mit tiefenpsychologisch fundierter Psychotherapie behandelt werden, dient in der PAL als Kontrollgruppe. Das zentrale Unterscheidungsmerkmal für die Zuordnung zu dieser Gruppe ist das Setting: Tiefenpsychologisch fundierte Psychotherapie wird definiert als einmal wöchentliche Therapie im Sitzen. Das Untersuchungsdesign für diese zweite Gruppe ist identisch mit dem für die Psychoanalysen. Das PAL-Design wird in identischer Form sowohl in Heidelberg als auch in Berlin umgesetzt, um eine möglichst hohe Fallzahl zu erreichen. Dabei wird angestrebt, insgesamt 30 Psychoanalysen mit 30 Psychotherapien vergleichen zu können.

Zu Beginn der vorliegenden Untersuchung befand sich der Datenpool noch in der Entstehungsphase. Von einigen Patienten existierten bereits Daten zu mehreren Zeitpunkten, andere begannen ihre Therapie erst. Es musste also ein Kompromiss zwischen mehreren Zielsetzungen gefunden werden: Erstens wollten wir eine ausreichend große Stichprobe untersuchen, um zumindest grobe inferenzstatistische Auswertungen zu ermöglichen. Zweitens sollte aber für jeden Patienten eine individuelle Verlaufsdarstellung möglich sein, durch Verwendung von zumindest mehr als zwei Messzeitpunkten. Dem Anliegen, möglichst viele Patienten zu möglichst vielen Zeitpunkten zu untersuchen, wurden wiederum durch zwei Faktoren Grenzen gesetzt: Zum einen standen, wie erwähnt, zunächst nur von einem Teil der Patienten Daten über einen längeren Behandlungszeitraum zur Verfügung – auch unter Berücksichtigung der voraussichtlichen Erhöhung dieser Anzahl im Verlauf der Erhebung. Zum anderen sollte der Arbeitsaufwand innerhalb eines angemessenen Rahmens für ein Promotionsvorhaben bleiben: Die Auswertung eines einzelnen Videobandes nahm durchschnittlich etwa 7-8 Stunden in Anspruch. Bei einer von uns gesetzten Obergrenze von etwa 500 Stunden Auswertungszeit ergab sich so eine Zahl von etwa 70 Videobändern, 30 in der Vor- und 40 in der Hauptstudie.

Für die Hauptstudie entschieden wir uns schließlich, zehn Patienten auszuwählen, von denen je vier Messzeitpunkte ausgewertet werden sollten: Zu Therapiebeginn (T1), nach sechs Monaten (T3), nach anderthalb Jahren (T7), und nach zwei Jahren (T9). Die von uns untersuchten zehn Fälle sind diejenigen, die als erste eine Behandlungsdauer von mindestens zwei Jahren erreichten, so dass die Videobänder zu allen oben genannten Zeitpunkten vorhanden waren. Da dies das einzige Einschlusskriterium darstellte, handelt es sich also um eine *unausgelesene* Gelegenheitsstichprobe aus der Gesamtgruppe der PAL-Patienten. Tabelle 1 stellt die wichtigsten Daten zur Beschreibung dieser Stichprobe dar.

Die Diagnosen wurden von den behandelnden Analytikern zu Beginn der Therapie gestellt. Aufgrund des teilweise unorthodoxen Gebrauchs der ICD-10 Kennziffern durch jene werden hier jedoch nur ihre verbalen Beschreibungen der einzelnen Symptomkomplexe wiedergegeben. Weil unsere Hypothesen (s.u.) störungsunspezifisch formuliert waren, schenkten wir den Diagnosen der Patienten im Rahmen unserer Auswertungen keine Beachtung (sie waren den Ratern auch bis zum Abschluss der Codierung nicht bekannt). Sie werden hier lediglich dargestellt, um die Vielfalt der klinischen Bilder in unserer Stichprobe zu illustrieren.

**Tabelle 1: Merkmale der von uns untersuchten Stichprobe**

Nummer	Geschlecht	Alter	Diagnose(n)
201	Weiblich	29	Depressiv-zwanghafte Persönlichkeit, Arbeitsstörung, Phobische Entwicklung in Sozialbeziehungen, Asthma Bronchiale
202	Weiblich	54	Mittelgradige depressive Episode, Ängste und depressive Störung gemischt
203	Männlich	33	Zwangsgedanken, Impotenz, herzneurotische Beschwerden, Durchfälle
204	Männlich	20	Schwere depressive Episode ohne psychotische Symptome, Vegetative Polysymptomatik
208	Männlich	48	Narzisstische Persönlichkeitsstörung, mittelgradige depressive Episode, Kopfschmerzsymptomatik, Morbus Reiter, Arterielle Hypertonie
209	Weiblich	33	Mittelgradige depressive Episode mit somatischen Symptomen, Zwangsstörung (Grübelzwang), Sexuelle Aversion und mangelnde sexuelle Befriedigung
214	Männlich	24	Zwangsstörung, vorwiegend Zwangshandlungen
215	Weiblich	41	Mittelgradige depressive Episode mit somatischen Symptomen
216	Weiblich	34	Generalisierte Angststörung, Nicht-organische Schlafstörungen, Leichte depressive Episode mit somatischen Symptomen, Spezifische isolierte Phobie, Abhängige Persönlichkeitsstörung
217	Weiblich	35	Hypomanie (als Abwehr), Störung der sexuellen Entwicklung, Schizoide Persönlichkeitszüge

Die Patienten verteilten sich ungleichmäßig auf die beiden Interviewer: Die weibliche Interviewerin führte die Gespräche mit zwei Männern und einer Frau, der männliche Interviewer mit den übrigen zwei Männern und fünf Frauen. Unserem Eindruck nach hatte die jeweilige Kombination von Patienten- und Interviewer-Geschlecht allenfalls einen marginalen Effekt auf die Daten.

Alle von uns untersuchten Patienten gehörten zur Gruppe der Psychoanalysepatienten. Ein Vergleich mit der Gruppe der tiefenpsychologisch fundiert behandelten Patienten erschien uns in Anbetracht der insgesamt sehr geringen Zahl verwendbarer Bänder wenig erfolgversprechend, weil ein solcher sehr große Effektstärken voraussetzen würde, um signifikante Unterschiede nachweisen zu können. Stattdessen sollten unterschiedliche Entwicklungen *innerhalb* der Psychoanalysegruppe betrachtet werden. Das Hauptaugenmerk legten wir auf die unterschiedliche Entwicklung der Beziehungsschemata von Patienten, deren Analysen als erfolgreich beurteilt werden können, im Vergleich mit solchen, in deren Behandlungen sich wenig positive Veränderungen zeigen.

Zur Überprüfung der Reliabilität unserer Messverfahren wurden im Rahmen der Vorstudie (Stadler, 2001) zehn Videobänder komplett von zwei unabhängigen Ratern (Daniel Leising und Karoline Stadler) durchcodiert. Dies erlaubte eine Bestimmung der Interater-Reliabilität. Acht dieser zehn Bänder waren identisch mit denen, die später im Rahmen der Hauptstudie als Zeitpunkt T9 noch einmal bearbeitet wurden (eine hundertprozentige Überschneidung war nicht zu erzielen, da zwei der zehn Bänder zu T9 erst im Laufe der Hauptstudie aufgenommen wurden). Um die Stichprobe für die Reliabilitätsermittlung möglichst groß werden zu lassen, wurden auch die zwei Bänder, die nicht Teil der Hauptstudie waren, zu diesem Zeitpunkt noch einmal gerated. Da der Rater, der die Bänder der Hauptstudie bearbeitete (D.L.) identisch mit einem der beiden Rater aus der Vorstudie war, ließ sich für diesen ein Index der Retest-Reliabilität errechnen. Das Intervall zwischen den beiden Beurteilungsdurchgängen belief sich auf etwa sechs Monate. Auch der Vergleich der Codierungen des ersten Raters (D.L.) im Rahmen der Hauptstudie mit denen des zweiten Raters (K.S.) aus der Vorstudie wurde durchgeführt. Er soll im Folgenden als *Cross-Reliabilität* bezeichnet werden, da hier sowohl unterschiedliche Rater als auch unterschiedliche Beurteilungszeitpunkte eingingen.

### **4.3 Messinstrumente**

Die Erfassung der Beziehungsnarrative erfolgte auf der Grundlage der Äußerungen der Patienten in den videographierten klinischen Interviews. Dem psychodynamischen Theoriehintergrund der PAL-Studie entsprechend waren die Interviewer gehalten, in den Gesprächen vor allem auf die Beziehungsgestaltung der Patienten sowie deren affektives Erleben einzugehen. Das Bild, das die Patienten dabei von ihren Interaktionen mit Anderen entwarfen, sollte in den Bereichen interpersonelles Verhalten und affektives Erleben durch unsere Messverfahren möglichst vollständig abgebildet werden.

Mit Crits-Christoph et al. (1994) halten wir die Beschränkung auf ein einziges, zentrales Beziehungsthema (wie im CCRT) für eine unzulässige Reduktion der Variabilität in den Beziehungsschilderungen. Analog dem QUAIN-ANSatz legten wir unsere Untersuchung daher so an, dass sich *Häufigkeitsprofile* von Beziehungsthemen ergeben konnten, aus denen sich die relative Bedeutsamkeit aller Beziehungsthemen zueinander ablesen lässt. Ein solches Vorgehen hat unter anderem den Vorteil, dass sich ein exakter Index der Stereotypizität von Beziehungsschilderungen (s.u.) ermitteln lässt, in den *alle* codierten Patientenäußerungen eingehen.

Die Prozedur zur Gewinnung der Daten sollte möglichst ökonomisch sein, um die Untersuchung einer ausreichend großen Stichprobe zu ermöglichen. Unter diesen Prämissen entschieden wir uns für folgendes Vorgehen:

Anstatt – wie üblich – die Berichte der Patienten zunächst zu transkribieren, entschlossen wir uns, die Kodierung *direkt vom Bildschirm* vorzunehmen, nachdem erste eigene Versuche mit dieser Methodik ermutigende Resultate gebracht hatten und eine Arbeit von Zander, Strack, Cierpka, Reich & Staats (1995) dies ebenfalls zulässig erscheinen ließ. Möglicherweise erhöht diese Vorgehensweise sogar die klinische Validität der Codierungen, da die Rater zusätzlich zur rein semantischen Information der gesprochenen Worte auch auf die paraverbale und mimische Information der Patienten zurückgreifen können.

Ein häufiges Problem in der Vermessung von Beziehungsrepräsentationen ergibt sich aus der Bestimmung der zu codierenden Einheiten: Der Rater muss wissen, *welche* Äußerung er codieren soll, bevor er sich für diese um eine inhaltliche Zuordnung bemühen kann. Auf den

Beziehungsepisoden-Ansatz und die mögliche Verwendung von Thought Units wurde bereits eingegangen (Crits-Christoph, 1998; Stinson, Milbrath, Reidbord & Bucci, 1994). Der Aufwand, den dieses Vorgehen in der Regel bedeutet, schien uns jedoch in keinem angemessenen Verhältnis zum erwartbaren Ergebnis zu stehen, so dass wir uns für eine eher pragmatische Lösung entschieden, indem wir die zwei – üblicherweise separaten – Arbeitsgänge zu einem einzigen zusammenfassten: *Jede Minute* der Interviews sollte codiert werden, wobei es dem Rater überlassen blieb, ob, und wenn ja, welche Äußerungen innerhalb der Minute er codieren wollte, allerdings mit der Einschränkung, dass jede Kategorie pro Minute nur einmal verwendet werden durfte. Der im Video eingeblendete Timer zeigte dabei den Beginn einer neuen codierbaren Einheit an. Das von Luborsky (1990a) geforderte *Completeness-Rating*, das die inhaltliche Beurteilbarkeit eines Textabschnittes beschreibt, ist in unserem Vorgehen implizit enthalten, da die Rater nur solche Beziehungsaussagen kategorisieren können, die im Hinblick auf das Beziehungserleben der Patienten ausreichend informativ sind. Auch auf die im CCRT übliche Bestimmung des Hauptobjektes verzichteten wir: Anstatt a priori festzulegen, mit welchem der Objekte der Patient in dem jeweiligen Abschnitt seines Berichts vor allem interagiert, sollte einfach das *erste* Auftreten jeder Kategorie pro Minute festgehalten werden.

Es sei betont, dass mit der Entscheidung gegen die Identifikation diskreter Beziehungsepisoden auch die Bindung der Methode an das Konzept interpersonal-affektiver Schemata teilweise relativiert wird: Die Theorie beinhaltet die Vorstellung, dass sich im Interaktionsverhalten eines Patienten bestimmte, charakteristische *Sequenzen* von Motiven, Emotionen und Verhaltensweisen auffinden lassen. Um solche zu identifizieren, wäre es nötig den Patienten dazu anzuhalten, möglichst viele interpersonelle Erlebnisse möglichst detailliert zu berichten, wie dies im Beziehungsepisoden-Interview (Luborsky, 1990b; Dahlbender, Torres, Reichert, Stübner, Frevert & Kächele, 1993) geschieht. Bei dem von uns untersuchten Videomaterial handelte es sich aber um relativ unstandardisierte OPD-Interviews (Schauenburg, Janssen & Buchheim, 1998). Die Anzahl der erzählten Beziehungsepisoden blieb letztlich dem Patienten überlassen. Manche berichteten im Verlauf einer Stunde über zwanzig solcher Geschichten, andere keine einzige – nach Stasch (pers. Mitteilung) ein häufiger Befund. Da wir auf die Gesprächsverläufe im Nachhinein keinen Einfluss mehr nehmen konnten, wäre die Identifikation vollständiger Beziehungsepisoden eine zu strenge Vorbedingung gewesen, die den Pool der verwendbaren Videos erheblich reduziert hätte. Die zeitlich-logische Verknüpfung der einzelnen Schema-Komponenten wird jedoch auch im CCRT nicht eingehalten. Dort werden die jeweils häufigsten Codierungen in den Bereichen W, RO und RS am Ende zu „dem“ CCRT zusammengefasst, *unabhängig davon, ob sie in den Erzählungen des Patienten jemals gemeinsam aufgetreten sind*. Das CCRT besteht schlicht aus den Modalwerten dreier Häufigkeitsprofile. Unsere Vorgehensweise folgt derselben Logik, mit dem Unterschied, dass wir die *vollständigen* Profile zur Grundlage unserer Berechnungen machten. Eine überarbeitete Version des CCRT, die die Verknüpfung der Komponenten untereinander berücksichtigt, für die statistische Testung jedoch viel größere Fallzahlen benötigt, wurde von Dahlbender, Albani, Pokorny & Kächele (1998) vorgestellt.

Innerhalb der einzelnen Minuten sollte *jede* Aussage, die ausreichend differenzierte Informationen über das Beziehungserleben in den relevanten Bereichen enthielt, codiert werden. Dazu sollte jeweils angegeben werden, welche inhaltliche Verhaltens- oder Affektqualität beschrieben wird, und ob der Patient diese sich Selbst (S) oder dem Objekt (O) zuschreibt. Die inhaltliche Kodierung erfolgte mittels SASB und der Klinischen Emotionsliste (s.u.). Dabei war es unerheblich, ob die Aussage eine konkrete Interaktion beschrieb

(“Gestern hat meine Schwester mir gesagt, ich solle die Spülmaschine einräumen”), eine wiederkehrende Konstellation (“Meine Schwester versucht andauernd, mir Vorschriften zu machen”) oder eine generelle Dimension des Beziehungsverhaltens (“Mir gegenüber ist meine Schwester total dominant”). Alle drei Aussagen müssten inhaltlich derselben SASB-Kategorie (1-5 O) zugeordnet werden und genauso wurde auch verfahren. Das soll nicht heißen, dass die erste genannte Aussage genau dasselbe bedeutet wie die letzte. Für die von uns angestrebten Profile von Beziehungsqualitäten kann die Frage, ob der Patient gehäuft Aussagen vom ersten, zweiten oder dritten Typ macht, jedoch vernachlässigt werden. Die Schlussfolgerung wäre dieselbe. Nur solche Interaktionen, an denen der Patient *selbst unmittelbar beteiligt* war, sollten codiert werden. Berichtete der Patient hingegen über Ereignisse, die sich zwischen zwei oder mehreren anderen Personen abspielten, ohne, dass er selbst involviert gewesen wäre, wurden keine Codes vergeben.

#### 4.3.1 Operationalisierung von Beziehungsschemata: Verhaltensebene

Für die Erfassung der von den Patienten beschriebenen interaktiven Handlungsweisen des Selbst und der Objekte griffen wir auf das oben beschriebene Codiersystem SASB (Benjamin, 1974) in der Clusterversion (Benjamin, 1984; Benjamin, 1986b) zurück. Dieses Verfahren vereint mehrere Vorzüge im Sinne unserer Zielsetzungen: Es bleibt sehr nah am beobachtbaren Verhalten, setzt keinen tiefenpsychologischen Theoriehintergrund voraus, und kann deshalb relativ problemlos Therapieschulen übergreifend eingesetzt werden. Darüber hinaus lässt es sich recht einfach erlernen. Nicht zuletzt wegen dieser Vorzüge gehört es mit Sicherheit zu den – auch in Deutschland – am häufigsten eingesetzten Verfahren zur Erfassung interpersonellen Verhaltens in der klinisch psychologischen Forschung (z.B. Fiedler, Backenstraß, Kronmüller & Mundt, 1998; Fiedler, Leeb, Ernst, Kohlhoff & Mundt, 1994; Fiedler, Vogt, Rogge & Schulte, 1994; Grande, 1997). Das Training für die Rater erfolgte im Selbststudium anhand des Manuals von Grawe-Gerber & Benjamin (1989) und der Arbeit von Grande (1997), sowie in häufiger persönlicher Rücksprache mit letzterem.

#### 4.3.2 Operationalisierung von Beziehungsschemata: Affektebene

Wir hatten uns vorgenommen, das selbstberichtete affektive Erleben der Patienten in ihren Beziehungen zu Anderen möglichst vollständig zu erfassen. Da dieser Versuch in Anbindung an den aktuellen Forschungsstand zum sprachlichen Ausdruck von Emotionen stattfinden sollte, orientierten wir uns zunächst an bereits existierenden Listen von Emotionsbegriffen wie der EMO-16 von Schmidt-Atzert & Hüppe (1996) oder der Aufzählung von Ulich & Mayring (1992). Anhand dieser kodierten wir versuchsweise die Äußerungen einiger Patienten. Recht schnell stellte sich jedoch heraus, dass in der von uns untersuchten klinischen Klientel eine ganze Reihe emotionaler Zustände immer wieder beschrieben wurde, die sich in den vorgegebenen Kategorien nicht sinnvoll unterbringen ließen. Dabei handelte es sich um solche Begriffe wie Hilflosigkeit, Verzweiflung, Niedergeschlagenheit, Erschütterung, Enttäuschung, Einsamkeit, Erschöpfung, aber auch einige positive Affektqualitäten wie Entspannung und Erleichterung. Wir entschlossen uns daher, zusätzliche Kategorien in unsere Affektliste aufzunehmen.

Im Rahmen der Vorstudie (Stadler, 2001) wurde das Instrument an 20 Videobändern weiterentwickelt, immer wieder modifiziert, neue Kategorien formuliert, und alte zusammengefasst. Es entstand schließlich eine Liste von 48 affektiven Qualitäten, die wir *Klinische Emotionsliste* (KE) nannten, und die wir für geeignet hielten, das Emotionserleben von psychisch

Gesunden und Kranken hinreichend vollständig abzubilden. Zum Abschluss der Vorstudie setzten wir die KE in der fertigen Form an 10 weiteren Bändern, unserer Reliabilitätsstichprobe (s.o.), ein, um die Reproduzierbarkeit der damit gewonnenen Daten zu überprüfen. Das Manual zur KE befindet sich im Anhang.

Die Zuordnung der Items zu den Gruppen der positiv oder negativ valenten Emotionen erfolgte nicht anhand der dem jeweiligen Affektzustand zugeschriebenen hedonischen Qualität – also der Frage, ob sich die Emotion „gut anfühlt“ – sondern danach, ob das *Objekt*, auf das sich die Emotion richtet, durch den emotionalen Zustand eine positive Bewertung erfährt. In den meisten Fällen führen die beiden erwähnten Zuordnungsregeln allerdings ohnehin zu identischen Ergebnissen: Nur „Sehnsucht“, die wahrscheinlich oft eher als etwas schmerzliches, daher tendenziell unangenehmes, empfunden wird, landet aufgrund der erwähnten Regel in der positiven Gruppe, da man sich nach etwas oder jemandem sehnt, das oder den man positiv bewertet. Zur Ökonomisierung der Codierung wurden die einzelnen Kategorien durch ihre Valenz, gefolgt von einer Kenn-Nummer abgekürzt, „Stolz“ also beispielsweise durch „P6“.

Man mag einwenden, dass das von uns entwickelte Kategoriensystem eigentlich zu differenziert ist. Die Anzahl von 48 Kategorien ist – verglichen mit anderen existierenden Emotionslisten, wie der EMO-16 oder der DAS (Merten & Krause, 1993) – tatsächlich recht hoch. Dennoch gehen wir davon aus, dass alle von uns für relevant erachteten Klassen sich, dem Ansatz von Mees (1985) entsprechend, aufgrund der dahinterstehenden semantischen Konstruktionen gut voneinander unterscheiden lassen. „Liebe“ ist nicht dasselbe wie „Zuneigung“, und beides ist nicht identisch mit dem Gefühl von „Nähe“. Es spricht jedoch nichts dagegen, diese und andere Begriffe aufgrund ihrer semantischen Ähnlichkeit zu übergeordneten Klassen zusammenzufassen, wenn man das System – zum Beispiel im Rahmen eines Folgeprojektes – vereinfachen will. Die 48 Items stellen unserer Ansicht nach eine *Obergrenze* der Differenzierung im Bereich des selbstberichteten affektiven Erlebens dar. Will man, wie in der vorliegenden Studie, unter anderem untersuchen, wie sich die emotionale Differenzierungsfähigkeit der Patienten im Therapieverlauf verändert (s.u.), sollte man diese jedoch nicht vorschnell durch eine zu geringe Itemzahl nach oben begrenzen.

### 4.3.3 Operationalisierung des Therapieerfolges

Als Therapieerfolgsmaße standen in der PAL gleich mehrere Datensätze zur Verfügung. Wir entschieden uns, aus den vorhandenen Selbstbeurteilungsmaßen diejenigen auszuwählen, die in der internationalen Literatur zur Psychotherapieforschung am weitesten verbreitet sind: Die Symptomcheckliste SCL 90-R (Franke, 1995) zur Selbstbeschreibung des aktuellen Symptomlevels durch die Patienten, und den IIP-D (Horowitz, Strauß & Kordy, 1994) als Index der selbstwahrgenommenen Belastung im interpersonellen Bereich. Für die Fremdbeurteilung des Therapieerfolges griffen wir auf die Daten der Heidelberger Umstrukturierungsskala (Rudolf, Grande & Oberbracht, 2000) zurück, auf der zwei Rater für jeden Patienten zu jedem Messzeitpunkt ein Konsensrating abgegeben hatten. Das Verfahren beschreibt die Durcharbeitung vorab als relevant gekennzeichnete psychodynamischer Behandlungsfoki im Therapieverlauf durch sieben qualitativ unterschiedliche Stadien: Nichtwahrnehmung des Fokus, Ungewollte Beschäftigung mit dem Fokusbereich, Vage Fokuswahrnehmung, Anerkennung und Erkundung des Fokusbereichs, Auflösung alter Strukturen im Fokusbereich, Neuordnung des Fokusbereichs, und Auflösung des Fokus.

Grundlage für die Einschätzungen der Umstrukturierung waren dieselben Interviews, anhand derer wir auch die Beziehungsrepräsentationen der Patienten analysierten. Bei diesem letzten der drei Erfolgsmaße lagen zum Zeitpunkt unserer Berechnungen zu zwei Patienten noch keine Daten vor, so dass hier nur eine Stichprobe von 8 Patienten untersucht werden konnte. Bei den Selbstbeurteilungsmaßen war hingegen N=10. Bei SCL 90-R und IIP-D verwendeten wir nur den jeweiligen Gesamtwert der Belastungsschwere, bei der HSCS den durchschnittlichen Score der Durcharbeitung über mehrere Behandlungsfoki hinweg.

#### **4.4 Zusatzcodierungen**

In unserer Vorstudie (Stadler, 2000) hatten wir vier Arten von Zusatzcodierungen ausprobiert, die verschiedene Aussagearten voneinander unterscheidbar machen sollten. Die einfachste Aussageform ist die Tatsachenbehauptung, z.B. "Ich lasse Andere lieber nicht so nahe an mich heran" (2-8 S). Solche Aussagen stellen die Regel dar und werden daher nicht mit einem Zusatzcode versehen. Demgegenüber bezeichnet

**V** Aussagen über Vergangenes, das heute nicht mehr gilt:

"Früher hab ich immer das gemacht, was die Anderen von mir verlangt haben, aber das hat sich total geändert." (2-5 S V)

**I** Aussagen des Interviewers, die vom Patienten bejaht werden:

Interviewer: "Sie achten sehr darauf, unabhängig zu bleiben, oder?"

Patient: "Ja, genau." (2-1 S I)

**K** Kommentare, die der Patient im Interview über seine Objekte macht:

"Ich finde das eigentlich ziemlich bewundernswert, was der X da gemacht hat." (1-2 S K)

**P** Phantasien des Patienten

"Also da hätte ich ihm am liebsten einfach eine reingehauen." (1-7 S P)

*Verneinungen* sollten zunächst als schlichte Aussagen darüber verstanden werden, dass Beschreibung X nicht auf die Realität zutrifft: "Also, dass ich da jetzt besonders anhänglich wäre kann man eigentlich nicht sagen" (2-4 S wird verneint). Eine solche Aussage rechtfertigt keine eigene inhaltliche Codierung. Anders verhält es sich jedoch, wenn der Patient bestimmte Erlebens- oder Verhaltensweisen vehement bestreitet. In solchen Fällen kann die Heftigkeit der Ablehnung einer Aussage die Codierung des genauen Gegenteils begründen:

Patient: "In meinem ganzen Leben hat sich noch nie irgend jemand um mich gekümmert."

(1-4 O wird so deutlich verneint, dass eine 1-8 O angebracht erscheint.)

Ebenso verhält es sich mit offensichtlich *ironisch* gemeinten Aussagen, die den Inhalt einer Aussage – oft durch paraverbale Mittel – ins Gegenteil verkehren. Die Grenze zwischen den beiden Bereichen ist hier sicherlich fließend.

Explizit ausgeschlossen wurden solche Interaktionen, die Patienten zwischen sich und dem *Interviewer* beschrieben:

Patient: “Ich hatte letztes Mal hier bei Ihnen das Gefühl, ich gehe Ihnen auf die Nerven”.

Dies deshalb, weil es sich bei der Untersuchungssituation um eine hochgradig künstliche Anordnung handelte, zum Beispiel dadurch, dass es nur alle sechs Monate zu einem einzelnen Kontakt kam. Die Repräsentativität dieser Interaktionen für das “typische” Beziehungserleben der Patienten erschien uns fraglich – obwohl sich bei einigen wenigen Patienten durchaus so etwas wie Übertragungskonstellationen mit den Interviewern entwickelten. Äußerungen wie die angegebene waren aber ohnehin ausgesprochen selten. Beschriebene Interaktionen mit dem behandelnden Therapeuten wurden genauso codiert wie die mit allen anderen Objekten.

## **4.5 Hypothesen**

### **4.5.1 Interaktives Verhalten des Selbst**

Benjamins (1993) Überlegungen lassen sich dahingehend operationalisieren, dass Patienten, die von ihrer Therapie profitieren, häufiger diejenigen Verhaltensweisen an den Tag legen sollten, die bei „psychisch Gesunden“ ihrer Meinung nach überwiegen, nämlich solche, die durch ein mittleres Ausmaß an Interdependenz und eine hohe Affiliation gekennzeichnet sind (vgl. 1.8). Mathematisch ausgedrückt bedeutet dies:

- Eine Verbesserung in den Therapieerfolgsmaßen geht für alle drei SASB-Foki mit einer Zunahme des Anteils der Cluster 2, 3 und 4 am gesamten berichteten eigenen Verhalten einher

Mit Crits-Christoph et al. (1994) gehen wir davon aus, dass psychisch gesündere Menschen sich durch variabelere Möglichkeiten, mit Anderen in Beziehung zu treten, auszeichnen. Entsprechend vermuten wir, dass Patienten, deren Therapien erfolgreich verlaufen, eine *Zunahme der Variabilität* ihres selbstberichteten Beziehungsverhaltens über die Zeit zeigen. Die Variabilität operationalisierten wir durch das von Strack (1993) entwickelte Dispersionsmaß, dessen Berechnung wir weiter unten kurz darstellen.

- Eine Verbesserung in den Therapieerfolgsmaßen geht mit einer Zunahme der Dispersion des eigenen Verhaltensprofils einher

### **4.5.2 Interaktives Verhalten der Objekte**

Für die Entwicklung des wahrgenommenen Objektverhaltens im Laufe erfolgreicher Therapien lassen sich weder bei Benjamin (1993) noch bei Crits-Christoph et al. (1994) Aussagen finden, die die Formulierung exakter Hypothesen gestatten würden. Der Einfachheit halber gehen wir hier von analogen Entwicklungen aus wie beim eigenen Verhalten, was sich auch gut begründen lässt: Patienten, die im Rahmen der Therapie gesünder werden, sollten in die Lage versetzt werden, sich mehr als früher solche Kontakte zu suchen, die ihnen

gut tun. Die Objekte müssten dann stärker als bestätigend, freundlich, oder hilfreich erlebt werden, aber auch als vertrauensvoller und offener. Auch eine Schärfung der Wahrnehmung für solche Aspekte interaktiven Verhaltens könnte zu einem Ansteigen des Anteils entsprechender Kategorien im Profil des Objektverhaltens beitragen. Für die Entwicklung des *introjektiven* Verhalten der Objekte hingegen lässt sich unseres Erachtens keine inhaltlich schlüssige Hypothese formulieren. Da dieses für den Patienten selbst auch von nachgeordneter Bedeutung ist, beziehen sich unsere Hypothesen in diesem Bereich daher nur auf die SASB-Foki I und II.

- Eine Verbesserung in den Therapieerfolgsmaßen geht für die SASB-Foki I und II mit einer Zunahme des Anteils der Cluster 2, 3 und 4 am gesamten berichteten Verhalten der Objekte einher

Eine zunehmend funktionale Wahrnehmung anderer Menschen sollte die Fähigkeit beinhalten, Andere differenzierter, situationsabhängiger zu erleben, anstatt ihr Verhalten anhand einiger weniger, immer gleicher Schablonen zu kategorisieren. Die Variabilität im Profil des Objektverhaltens sollte daher, ebenso wie die des eigenen Verhaltens, mit dem Erfolg der Therapie kovariieren. Wiederum beschränken wir diese Hypothese auf die SASB-Foki I und II, also diejenigen Verhaltensweisen, von denen der berichtende Patient unmittelbar betroffen ist.

- Eine Verbesserung in den Therapieerfolgsmaßen geht mit einer Zunahme der Dispersion im Profil des Objektverhaltens einher

#### 4.5.3 Eigene Emotionen

Entsprechend der Annahme einer zunehmenden Affiliation, also positiven Bezogenheit, im eigenen Verhalten der Patienten, sollte sich auch der Anteil solcher Emotionen, in deren semantischer Struktur eine positive Selbst- oder Objektbewertung enthalten ist, am eigenen Emotionsprofil erhöhen. Um die Logik dieser Hypothese möglichst klar umzusetzen, verwendeten wir hier den Quotienten aus *Anzahl positiver Emotionen geteilt durch Anzahl negativer Emotionen* und ließen die sonstigen Emotionen für die Berechnung außer Acht. Aufgrund des ausgesprochen geringen Anteils, den diese letzteren an den Profilen der meisten Patienten ausmachten (s. Ergebnisteil), wären aber auch bei Verwendung eines anderen Bruches (z.B. *Anzahl positiver Emotionen durch Anzahl aller übrigen Emotionen*) keine nennenswert anderen Ergebnisse zu erwarten gewesen.

- Eine Verbesserung in den Therapieerfolgsmaßen geht mit einer Zunahme des Quotienten *Anzahl positiver Emotionen durch Anzahl negativer Emotionen* einher

Entsprechend der Annahme in der OPD (Arbeitskreis OPD, 1996), dass neurotische Konflikte mit einem bevorzugten Erleben bestimmter Affektqualitäten einhergehen, nehmen wir an, dass sich die Gesundung eines Patienten auch in einer größeren Spannweite des affektiven Repertoires äußern sollte. Patienten, deren Behandlungen erfolgreicher verlaufen, sollten daher eine zunehmende Varianz ihres Emotionsprofils aufweisen – einzelne Affekte dürften verglichen mit anderen nicht mehr so sehr im Vordergrund stehen.

- Eine Verbesserung in den Therapieerfolgsmaßen geht mit einer Zunahme der Dispersion des eigenen Emotionsprofils einher

#### 4.6 Exkurs: Das Dispersionsmaß nach Strack (1993)

Die Variabilität des interpersonal-affektiven Repertoires einer Person lässt sich auf ganz unterschiedliche Art und Weise operationalisieren. Im Ansatz von Crits-Christoph & Luborsky (1990) wird die "Pervasiveness" des CCRT noch dadurch ausgedrückt, in wievielen der insgesamt berichteten Beziehungsepisoden Komponenten des zum CCRT bestimmten Musters vorkommen. Diesem Ansatz liegt jedoch die oben bereits problematisierte Setzung zugrunde, dass sich für jeden Patienten genau ein zentrales Beziehungsmuster identifizieren lässt. Ein wesentlich ausgefeilteres Maß ist demgegenüber das der Dispersion (DISP) von Strack (1993), da es sich nicht auf die Analyse einer einzelnen Verhaltens- oder Erlebens-kategorie beschränkt, sondern *alle* verfügbaren Daten einbezieht. Die Logik dieses Indexes soll im Folgenden kurz umrissen werden (eine gute Darstellung findet sich auch bei Cierpka et al., 1998).

Gesucht ist eine Formel zur Berechnung der Variabilität eines Nominaldatenprofils. Strack (1993) listet eine Reihe bereits existierender Lösungsvorschläge (z.B. Entropie, Redundanz) auf, befindet sie jedoch aus jeweils unterschiedlichen Gründen für unzureichend. Ihr eigener Ansatz beruht auf dem Konzentrationsmaß (Konz) von Gini, das sich wie folgt berechnen lässt:

$$\text{Konz} = 1 - \sum_{i=1}^k (p_i)^2$$

Dabei ist  $p_i$  die relative Häufigkeit, mit der jede Kategorie  $i$  besetzt wurde. Die Anzahl der benutzten Kategorien ist  $k$ . Da die Obergrenze der Konzentration jedoch auch von der Anzahl der zur Verfügung stehenden (benutzbaren) Kategorien  $K$ , sowie der Anzahl der zu kategorisierenden Objekte  $m$  abhängt, muss Konz noch in Relation zu der im vorliegenden Fall jeweils maximal erreichbaren Konzentration (KonzMax) gesetzt werden. Konzmax entspricht – bei gleichem  $K$  und  $m$  – dem Konzentrationswert bei einer möglichst gleichförmigen Verteilung der Objekte auf die Kategorien. Werden zum Beispiel 57 Objekte auf 10 Kategorien verteilt, ergibt sich die ausgeglichene Verteilung, wenn in 7 Kategorien je 6 Objekte landen und in den verbleibenden 3 Kategorien je 5 Objekte ( $7*6 + 3*5 = 57$ ). Man teilt  $m$  durch  $K$  ( $57$  durch  $10$ ) und erhält, weil nur ganzzahlige absolute Häufigkeiten zulässig sind, "5 Rest 7". Wir bezeichnen die erste Zahl, das ganzzahlige Divisionsergebnis von "m durch K", als  $a$ . Diese Zahl gibt an wieviele Objekte zunächst in jede der  $K$  Kategorien fallen müssen, damit sich eine Gleichverteilung ergibt. Nun muss aber auch noch der Rest (7), den wir  $b$  nennen wollen, möglichst gleichmäßig auf die Kategorien verteilt werden. Dies ist der Fall, wenn jedes der übriggebliebenen Objekte eine andere Kategorie besetzt, so dass in 7 Kategorien zu den 5 Objekten noch jeweils eins dazu kommt. Allgemein resultiert für die jeweils maximal ausgeglichene Verteilung

$$m = (K-b) * a + b * (a+1)$$

und

$$\text{KonzMax} = 1 - ((K-b) * (a/m)^2 + b * ((a+1)/m)^2)$$

Wie man sieht ergibt sich im Fall, dass die Anzahl der zu klassifizierenden Objekte kleiner als die Zahl der verfügbaren Kategorien ist, für a der Wert 0, so dass der erste Term in der obigen Gleichung ganz wegfällt.

Die Dispersion DISP berechnet sich nun als der Quotient

$$\text{DISP} = \text{Konz} / \text{KonzMax}$$

Dieser Wert liegt immer zwischen 0 und 1. Je höher DISP, desto gleichmäßiger verteilen sich die Objekte auf die Kategorien.

#### **4.7 Statistische Modellierung**

Wir entschieden uns, unsere Zusammenhangshypothesen anhand eines einfachen hierarchisch linearen Modells (Bryk & Raudenbush, 1987) zu überprüfen.

Als erstes wurden für jeden Patienten die individuellen Verläufe der Therapieerfolgsmaße und der oben genannten Anteils- und Dispersionswerte der Verhaltens- und Emotionsprofile über die Zeit nach dem Kleinste-Quadrate-Kriterium durch eine Gerade angenähert. Wir befanden eine solche Modellierung für hinreichend, weil unsere Hypothesen sich ausschließlich auf die Zu- oder Abnahme bestimmter Werte bezogen. Die Passungsgüte des linearen Modells wurde demgegenüber bewusst vernachlässigt: Ein symmetrisch U-förmiger Verlauf der Rohdaten über die Zeit lässt sich beispielsweise durch eine Gerade nur sehr ungenau, mit großen Residuen, beschreiben. Dennoch wäre ein linearer Fit für unsere Zielsetzung auch in diesem Fall die angemessene Lösung, da eine solche Kurve eben keine systematische Zu- oder Abnahme beschreibt – die resultierende Gerade hätte folgerichtig die Steigung 0.

Auf diese Weise ergeben sich für jeden Patienten Geradengleichungen, die die Entwicklung der Therapieerfolgsmaße und des Beziehungserlebens in den verschiedenen Bereichen jeweils durch einen einzigen Wert, die Steigung, beschreibbar machen. Die Achsenabschnitte, also die extrapolierten Ausgangsniveaus der Geraden wurden von uns nicht weiter untersucht.

Zur Überprüfung unserer Hypothesen errechneten wir nun in einem zweiten Schritt einfache Pearson-Korrelationen zwischen den Steigungsindizes. Ein solcher Korrelationskoeffizient beschreibt den *linearen Zusammenhang zweier linearer Trends*, also das Ausmaß in dem eine Zu- oder Abnahme im einen Messbereich mit einer Zu- oder Abnahme im anderen Messbereich einhergeht. Da wir ausnahmslos gerichtete Hypothesen formuliert hatten, berechneten wir für diese Korrelationen *einseitige* Signifikanztests, unter der Annahme, dass sich die Trends in der Population, aus der unsere Stichprobe stammt, normal verteilen. Auch bei denjenigen Signifikanzprüfungen, die nicht der Hypothesentestung im engeren Sinne dienen, verwendeten wir einseitige Tests, weil in jedem Fall die Richtung des erwarteten Effektes vorab formuliert werden konnte.

Der Einsatz eines HLM erschien auch deshalb angezeigt, weil sich dadurch eine Besonderheit unseres Datensatzes angemessen berücksichtigen ließ: Während nämlich für alle Indizes des Beziehungserlebens sowie für die IIP- und Umstrukturierungswerte Daten für die Zeitpunkte T1, T3, T7 und T9 vorhanden waren, existierten Daten der SCL 90-R nur für die

Messzeitpunkte T1, T5 und T9. Die Modellierung der Verläufe durch Geraden funktioniert jedoch unabhängig von der Anordnung der Messzeitpunkte. Ebenso ist es unproblematisch, die Steigungen von Geraden zu korrelieren, die auf der Grundlage unterschiedlicher Messzeitpunkte ermittelt wurden. Sämtliche Berechnungen wurden mit SPSS 10.0 durchgeführt.