

Kap. 4

4. Die Untersuchung eines Behandlungsverlaufes mit der maschinellen Inhaltsanalyse

4.1 Einleitende Bemerkungen zu der hier untersuchten psychoanalytischen Behandlung

4.2 Zur deskriptiven Statistik der Rohdaten

4.3 Die Vorhersage klinischer Konzepte durch die maschinelle Inhaltsanalyse

4.1 Einleitende Bemerkungen zu der hier untersuchten psychoanalytischen Behandlung

Die hier vorgelegten Untersuchungen wurden erst – wie das gesamte Forschungsprojekt, in dessen Rahmen diese Untersuchungen stehen – dadurch ermöglicht, dass seit 1969 ein Psychoanalytiker seine Behandlungen mit Tonbandaufnahmen systematisch und vollständig erfasste. Dies ist unseres Wissens die erste vollständig aufgezeichnete Psychoanalyse, über die wir im deutschen Sprachraum verfügen.

In den USA hatte sich die Methode der Tonbandaufzeichnung schon seit längerem als für Forschungszwecke unerlässliches Hilfsmittel durchgesetzt, so dass dort an verschiedenen Stellen umfangreiche Archive von Behandlungsprotokollen bestehen (s.d. LUBORSKY und SPENCE, 1971).

Die Aufzeichnung einer Behandlung mit Hilfe des Tonbandes führt zu besonderen Belastungen; weniger für den Patienten, wie dies lange genug von vielen engagierten Gegnern solcher Tonbandaufzeichnungen ins Feld geführt wurde, sondern weitaus mehr für den behandelnden Psychoanalytiker, der sich dadurch in einem sonst nicht möglichen Ausmaß vor der wissenschaftlichen Öffentlichkeit exponiert (s.d. GILL et al., 1968, SIMON et al., 1970).

Für die hier untersuchte Behandlung entstand eine weitere Belastung noch dadurch, dass die Existenz und Kontinuität des Forschungsprojektes an die Durchführung dieser Behandlung gebunden war, eben weil es noch keine abgeschlossenen aufgezeichneten Behandlungen beim Beginn des Projektes gab. Die ausführliche Darstellung der damit verbundenen behandelungstechnischen Probleme, speziell der Gegenübertragungsreaktionen des Analytikers wird an anderer Stelle gegeben werden (THOMÄ, 1976); als Einleitung zu den hier vorgelegten inhaltsanalytischen Untersuchungen muss eine kurze zusammenfassende Darstellung für die Einführung in die klinische Problematik des Patienten genügen. Im Rahmen des Teilprojektes „Systematische Beschreibung des Behandlungsprozesses“ haben wir bereits eine Darstellung des Behandlungsverlaufes gegeben, auf die wir hier nur verweisen können (THOMÄ, KÄCHELE und SCHAUMBURG, 1973).

Der Patient:

Es handelt sich bei dem Patienten um einen, bei Behandlungsbeginn 21-jährigen Studenten, der im Jahre 1969 wegen häufiger Anfälle von paroxysmaler Tachykardie und schweren Angstzuständen zur stationären Aufnahme kam. Die Angstsymptomatik hatte sich langsam entwickelt und führte zu einer zunehmenden Einschränkung der Mobilität des Patienten. Eine für den Patienten sehr kränkende Zurückweisung durch ein Mädchen führte dann zur Entwicklung der paroxysmalen Tachykardien. Außerdem bestanden Angstanfälle mit heftiger Übelkeit und Erbrechen; diese machten es dem Patienten unmöglich, allein das Haus zu verlassen. Hinzu kamen erhebliche Minderwertigkeitsgefühle, ein durchgängiges Gefühl der Unsicherheit und der Existenzangst. Diese Schwierigkeiten behinderten den Patienten so vollständig, dass er sein Studium unterbrechen musste.

Aus der Lebensgeschichte des Patienten können wir hier nur kurz die wichtigsten Angaben zusammenstellen:

In der Kindheit bis zum Schuleintritt bestanden zahlreiche starke infantile Ängste. Der Patient wuchs als ältester Sohn in Abwesenheit des Vaters auf und entwickelte eine sehr enge Bindung an seine Mutter, die sich wegen seiner Angst im Dunkeln oder beim Alleinsein im Laufe der Jahre verstärkte. Etwa im vierten Lebensjahr kam der Vater zurück. Im Zusammenhang

mit Träumen, in denen der Patient durch böse Blicke eines Mannes oder durch dessen körperliche Bedrohungen (mit einer Kneifzange) bestraft wurde, verstärkten sich die Angstzustände. Er ist der älteste Sohn einer überängstlichen und verwöhnenden Mutter und eines Vaters, der recht zurückgezogen seinen beruflichen Interessen nachstrebte und als Physiker erfolgreich ist. Nach dem Krieg war er mehrere Jahre im Ausland dienstverpflichtet und deshalb die meiste Zeit abwesend.

Die Verwöhnung durch die Mutter war an die Einhaltung ihrer Ideale durch den Patienten gebunden: So entwickelte sich ein überbraver, überangepasster Junge, der Neugierde und Aktivität extrem einengte, um sich die Liebe der Mutter zu erhalten. Sein Rivalisieren mit zwei jüngeren Brüdern wurde unterdrückt. Seine Leistungen in der Schule blieben hinter dem zurück, was er aufgrund seiner guten Intelligenz hätte leisten können. Eine Konzentrations- und Arbeitsstörung war dadurch verdeckt, dass der Patient jeweils mühelos das Klassenziel erreichte; solange er ohne Anstrengungen zum Ziel kam, fühlte er sich wohl. Schon die geringsten Belastungen führten zu Unlust bzw. Angst und erschütterten sein brüchiges Selbstgefühl; obwohl er sich durch seine Wohlerzogenheit überall beliebt machen konnte, erinnert der Patient keine Phase seines Lebens, in der er auf irgendeinem Gebiet mit sich selbst zufrieden war oder eine entsprechende Sicherheit aus einer zwischenmenschlichen Beziehung hätte ziehen können.

In der Zeit bis zum 12. Lebensjahr bestand eine ausgeprägte Schulphobie. Diese besserte sich unter der verwöhnenden Fürsorge der Mutter, die ihren Sohn während eines längeren Zeitraumes auf dem Schulweg begleitete und ihm die Schulaufgaben abnahm. Der Tod des geliebten Großvaters an Herzversagen war dann ein wichtiger Punkt in der Ausbildung der herzbezogenen Beschwerden. In der Pubertät folgte dann eine klinische Untersuchung und Behandlung des Patienten wegen „Herz-Kreislaufbeschwerden“. Die von ihm erinnerte Diagnose, „er habe ein schwaches Herz“, wurde mit der ärztlichen Empfehlung verbunden, dass der Junge geschont werden müsse. Damit ging eine kurze Phase sportlich-körperlicher Aktivität zu Ende mit Umschlag in Passivität und Abhängigkeit. Vom 12. Lebensjahr an war der Patient bei verschiedenen Ärzten wegen Angstgefühlen, Herzsensationen und Extrasystolen in Behandlung.

Während der Pubertät versuchte er sich mit starkem Willensaufwand von seiner Mutter zu befreien. Er erwies sich als unfähig, einen sportlichen Konkurrenzkampf bis zum Ende durchzustehen. Es gehörte zu einer seiner Verhaltensstereotypen, dass er kurz vor dem Sieg aufgeben musste. In bewusster Entschiedenheit versuchte der Patient die pruden Ideale seiner Mutter über Bord zu werfen, ohne dass er seiner Erfolge bei Mädchen froh werden konnte. Seine Freundschaften hatten narzisstischen Charakter. Sie brachten ihm ebenso wie die Masturbation mehr Schuldgefühle als Befriedigung ein.

Den letzten Rest von Selbstsicherheit verlor der Patient nach einer Zurückweisung durch ein Mädchen. Diese narzisstische Kränkung führte zur Dekompensation und zum Freiwerden frei flottierender Ängste, die fast andauernd als körperlicher Schaden erlebt wurden.

Wegen eines nächtlichen Angstanfalles, verbunden mit einer paroxysmalen Tachykardie erfolgte die Einweisung als Notfall in eine Innere Klinik, einige Monate vor der Überweisung zur Psychotherapie. Seit diesem Anfall diffuse Angstzustände mit Übelkeit und Brechreiz, die es dem Patienten unmöglich machten, allein auf die Straße zu gehen.

Diagnostisch handelt es sich um eine Angstneurose, deren Intensität – wobei besonders die Diffusität der angstausslösenden Situationen geradezu extreme Ausmaße angenommen hat – zusätzlich die Annahme einer charakterneurotischen Störung nahe legt. Das Ausmaß der Einschränkung machte die Durchführung einer psychoanalytischen Behandlung notwendig.

Der Behandlungsverlauf

Wegen der Schwere der Symptomatologie und aus äußeren Gründen musste der Patient zwei Jahre lang stationär behandelt werden. Im dritten Jahr konnte der Patient ambulant zur Behandlung kommen, musste aber immer noch das Taxi benützen, da er keine öffentlichen Verkehrsmittel ertragen konnte. In diesen drei Jahren konnte er keiner regelmäßigen körperlichen oder geistigen Tätigkeit nachgehen. Erst im vierten Jahr nahm er sein Studium an einer 100 km entfernten Universität wieder auf, was eine Verringerung der Behandlungsfrequenz auf zwei Wochenstunden notwendig machte. Bei einer sich ständig verbessernden sozialen und motorischen Entfaltung des Patienten blieben seine enormen Minderwertigkeitsgefühle lange Zeit unverändert. Die Behandlung führte nach der Durcharbeitung ödipalen Materials immer stärker zu einer Zentrierung auf eine symbiotische Problematik, in deren Mittelpunkt die Unlust und Unfähigkeit des Patienten stand, aus irgendeinem Tun oder Erleben Befriedigung oder Zufriedenheit zu ziehen. Die Bearbeitung dieser Selbstwert-Problematik, die mit der noch nicht vollzogenen Trennung aus der unbewussten Symbiose mit der Mutter zusammenhängt, beherrscht die Behandlung ca. seit der 400. Stunde. In der Zwischenzeit hat der Patient sein Studium mit glänzenden Examina abgeschlossen, bewährt sich als Referendar und ist hinsichtlich seiner Angstsymptomatik fast beschwerdefrei.

Der Teil der Behandlung, den wir bei unserer Untersuchung bearbeitet haben, bezieht sich auf den Zeitraum der Stunden 001 - 630. Die ausführlichere Darstellung des Behandlungsverlaufes, wobei besonders die Darstellung der schwierigen technischen Probleme im Mittelpunkt steht, ist zur Zeit in Vorbereitung (THOMÄ, 1976).

4.2 Zur deskriptiven Statistik der Rohdaten

4.2.1 Die statistischen Beschreibung der Rohdaten der Untersuchung

4.2.2 Mittelwert und Standardabweichung der Variablen

4.2.3 Die Variabilität der Kategorien im Behandlungsprozess

4.2.4 Die Ähnlichkeit des Sprachinhaltes bei Patient und Analytiker

4.2.1 Die statistische Beschreibung der Rohdaten der Untersuchung

Die maschinelle Inhaltsanalyse mit dem EVA-System lieferte für jede der 130 untersuchten Behandlungsstunden, die als „summarizing unit“ (AULD und MURRAY, 1955) zugrundegelegt waren, absolute und prozentuale Häufigkeitswerte der 61 Kategorien für den Patienten- und Analytikertext getrennt.

Als Ausgangswerte für die weiteren Berechnungen verwendeten wir die auf den Textumfang relativierten Häufigkeiten, die auch als KT-Werte (d.h. Kategorien pro Texteinheit) bezeichnet werden (DEICHSEL, 1975). Um eine Übersicht über diese von EVA erzeugten Daten zu gewinnen, sollen im folgenden einige statistisch deskriptive Angaben gemacht werden:

Dabei wird zunächst der Mittelwert und die Standardabweichung der erhaltenen Rohdaten beschrieben (4.2.2), der Variationskoeffizient der Variablen diskutiert (4.2.3) und dann die Ähnlichkeit des Sprachinhaltes von Patient und Analytiker über die 130 Behandlungsstunden hinweg untersucht werden (4.2.4).

4.2.2 Mittelwert und Standardabweichung der Variablen

Die folgende Tabellierung gibt den Mittelwert und die Standardabweichung für jede Variable über die untersuchte Stichprobe von 130 Stunden an. Dabei werden die Werte für die einzelnen Variablen für Patient und Analytiker parallel dargestellt, um einen Vergleich zu ermöglichen; außerdem trennen wir zur besseren Übersicht zwischen den Kategorien 1. Ordnung (Variable Nr. 1-45) und den Kategorien 2. Ordnung (Variable Nr. 46-61):

Tabelle: Mittelwert und Standardabweichung der Kategorien erster Ordnung

Kategorie	Patient		Analytiker	
	MEAN	S.D.	MEAN	S.D.
Male Role	16.0246	22.5693	16.5977	24.9721
Female Role	25.3623	32.8881	22.8131	33.3458
Neuter Role	14.4138	18.5518	9.4669	14.8387
Job Role	5.4208	10.6725	2.1154	5.4694
Large Group	9.2869	14.3638	0.0962	0.7724
Body Part	20.6208	18.7979	19.3300	28.1224
Food	0.0923	1.0525	0.7215	8.2268
Clothing	0.5815	2.7555	0.7877	4.2641
Tool	3.2446	7.4587	4.5769	9.8799
Natural Object	4.5515	8.5142	4.8277	15.8190
Non specific Object	50.9254	26.2775	30.2623	23.7962
Time Reference	70.4846	38.9780	44.8123	37.6626
Space Reference	20.6069	17.1370	28.1477	22.7830
Quantity Reference	11.3846	13.6036	10.1254	15.5253
Social Place	15.7100	16.2993	14.3808	19.5309
Natural World	5.4331	8.9532	6.1908	10.5525
Ideal World	8.7169	14.2312	18.3392	26.5993
Deviation	11.8469	17.0839	2.2415	8.9445
Action Norm	5.4115	9.6739	4.7408	8.5834
Message Form	23.3969	20.1253	37.8777	30.3077
Thought Form	54.1554	35.2664	74.2715	46.6148
Arousal	17.0469	17.0484	22.2123	25.7620
Urge	10.6338	18.4847	13.3877	18.7016
Affection	11.1108	15.3804	3.7292	8.5438
Pleasure	21.4300	31.2388	20.8369	43.8915
Distress	93.6554	57.6340	53.0015	35.8517
Anger	24.1246	35.5533	27.7523	45.9965
Sense	8.2146	13.6386	14.0846	16.4362
Think	25.9785	22.6140	23.3200	19.6656
If	4.3500	6.9856	1.2200	3.7208
Not	2.6808	5.3536	2.2277	5.1554
Defense Mechanism	1.9777	6.9418	2.2100	8.3357
Good	1.1231	3.8681	7.0931	17.3655
Bad	0.8185	6.0346	1.6485	8.4394
Communicate	5.5423	9.1180	9.0508	14.2057
Approach	3.3162	7.0489	4.1400	10.4821
Guide	3.1692	7.6736	2.7615	6.4550
Follow	1.2646	4.1413	1.9831	6.0201
Attack	6.1469	9.7835	16.2231	22.7567
Avoid	0.5954	2.4459	4.3508	11.1864
Attempt	8.3508	13.5714	6.1008	10.2662
Get	1.3615	3.6276	4.2008	12.5149
Expel	2.1600	8.3849	4.1831	8.3267
Work	9.5562	18.6069	4.8600	10.9183
Move	2.5715	6.4213	4.6492	10.3847

Abbildung 24 a

Tabelle: Mittelwert und Standardabweichung der Kategorien zweiter Ordnung

Kategorie	Patient		Analytiker	
	MEAN	S.D.	MEAN	S.D.
Academic	21.3400	25.0106	0.3754	2.2241
Community	12.9592	16.0350	0.8592	3.2269
Family	19.7815	32.8059	21.8969	35.2594
Medical	19.2038	26.4223	9.2569	18.3825
Higher Status	3.5615	8.9182	1.3654	4.4127
Peer Status	4.1838	8.2627	4.4123	12.3336
Lower Status	18.2962	29.9176	10.4054	22.1507
Overstate	2.9492	6.7459	1.6223	5.1225
Understate	5.5231	7.7719	3.3400	8.3394
Signstrong	6.0562	10.4075	14.1292	20.7856
Signweak	2.8769	7.4335	10.8508	19.0827
Signaccept	7.3777	12.1207	11.5692	19.7906
Sex Sense	30.7785	33.8271	40.0369	48.9361
Auth Theme	13.2585	18.4772	13.6108	21.0317
Danger Theme	1.0415	3.5784	1.0092	3.4828
You	21.6892	25.4132	0.5985	4.4583

Abbildung 24 b

Die tabellarische Übersicht (Abb. 24 a, S. 80 und 24 b, S. 81) zeigt, dass der mittlere Anteil der Kategorien am Text sehr unterschiedlich ist. Die Kategorie DISTRESS des Patienten weist z.B. einen Mittelwert von 93,66 Y auf, also knapp 10% im Mittel der 130 Stunden, während andere Kategorien, z.B. die Kategorie AVOID beim Patienten mit einem mittleren Anteil von 0,60Y zu den Kategorien gehört, die nur wenig auftreten. Ein Blick auf die Maximum-Minimum-Werte zeigt weiterhin, dass die Verteilung der meisten Kategorien extrem schief ist, was die Verwendung einer Vielzahl von statistischen Verfahren, die eine Normalverteilung voraussetzen, problematisch macht^{*)}. (s.d. KRIZ, 1975, S. 140 ff.).

Der Vergleich der Mittelwerte von Patient und Analytiker ergibt eine überraschend große Ähnlichkeit. Listet man die Kategorien nach ihrem prozentualen Anteil absteigend auf, so wird diese Ähnlichkeit der Gliederung des Sprachschatzes besonders deutlich: Abbildung 25, S. 82):

^{*)} Wir haben für andere Fragestellungen zu einem späteren Zeitpunkt eine Normalisierung der Daten durchgeführt, die jedoch nur unvollkommen die Verteilungseigenarten dieser Daten eliminieren konnte. Die Normalisierung wurde mit Hilfe des Programmes NORMA 3 von Herrn Dr. Timm, Universität Freiburg durchgeführt. Die Adaption des Programmes auf der TR440 am Rechenzentrum der Universität Ulm besorgte Dipl.-Inf. Mergenthaler.

Rang	Patient	Rang	Analytiker
1	DISTRESS	1	THOUGHT FORM
2	TIME REF.	2	DISTRESS
3	THOUGHT FORM	3	TIME REF.
4	NON SPEC. OBJECT	4	MESSAGE FORM
5	THINK	5	NON SPEC. OBJECT
6	FEMALE ROLE	6	SPACE REF.
7	ANGER	7	ANGER
8	MESSAGE FORM	8	THINK
9	PLEASURE	9	FEMALE ROLE
10	BODYPART	10	AROUSAL
11	SPACE REF.	11	PLEASURE
12	AROUSAL	12	BODYPART
13	MALE ROLE	13	IDEAL WORLD
14	SOCIAL PLACE	14	MALE ROLE
15	NEUTER ROLE	15	ATTACK
16	DEVIATION	16	SOCIAL PLACE
17	QUANTITY REF.	17	SENSE
18	AFFECTION	18	URGE
19	URGE	19	QUANTITY REF.
20	IDEAL WORLD	20	NEUTER ROLE
21	SENSE	21	AFFECTION
22	ATTACK	22	DEVIATION

Abbildung 25

Diese Übersicht über 22 der 45 Kategorien 1. Ordnung zeigt, dass diese Kategorien mit einem Textanteil, der größer als 10 Promille ist, eine sehr ähnliche Rangfolge aufweisen. Die Bestimmung der Rangkorrelation nach SPEARMAN über alle 45 Kategorien ergibt entsprechend auch einen Koeffizient von $r_s = + 0.85$, der für $N = 45$ hochsignifikant ist.

Die Angleichung der mittleren Häufigkeiten der Kategorien bei Patient und Analytiker stellt ein bisher in der Literatur in dieser Form noch nicht beschriebenes Phänomen dar. Zwar konnte LAFFAL (1967) nachweisen, dass sich bei einer Drei-Personen-Unterhaltung die Gesprächsinhalte der drei Gesprächspartner ähnlich sind; seine Untersuchungsmethode basierte jedoch auf der Verwendung der Rangkorrelationstechnik über alle Kategorien des inhaltsanalytischen Wörterbuches. Da nach unserer Erfahrung die Häufigkeitsverteilungen der einzelnen Wörter, wie in Kapitel 3 beschrieben, zu ähnlichen ZIPF-Verteilungen der Inhaltskategorien führen, ist der LAFFAL'sche Befund einer Ähnlichkeit zwischen allen Kategorien ein gutes Stück weit trivial. Hier beschreiben wir jedoch die Angleichung der Mittelwerte der einzelnen Kategorien als Ausdruck einer systemtheoretischen Eigenschaft des hier untersuchten Textkorpus. Eine detaillierte Aufweisung dieses Abstimmungsprozesses bedarf aber noch der genaueren Untersuchung des zeitlichen Ablaufes anhand der einzelnen Stunden. Aufgrund der Befunde der Wortschatzanalyse ist allerdings mit großer Sicherheit anzunehmen, dass es hier vorwiegend um die Abstimmung zwischen den häufigsten Wörtern, die bei dem Patient und dem Analytiker sehr oft vorkommen, geht. Wir werden später die Ähnlichkeit des Sprachinhaltes einzelner Stunden noch gesondert aufgreifen.

4.2.3 Die Variabilität der Kategorien im Behandlungsprozess

Für die Brauchbarkeit des Wörterbuches als Instrument der Verlaufsbeschreibung ist die Variabilität der Kategorien von entscheidender Bedeutung. Ein Instrument zur Messung der Veränderung muss sensibel für tatsächliche Veränderungen im Untersuchungsmaterial sein^{*)} – ein Umstand, der die Verwendung mancher Verfahren gerade im Bereich der Verlaufsforschung als problematisch erscheinen lässt.

CHASSAN (1967), der sich besonders ausgiebig mit den statistischen Problemen im Bereich der psychiatrischen und psychopharmakologischen Verlaufsuntersuchungen befasst hat, spricht in diesem Zusammenhang von der notwendigen „Irreliabilität“ der Messinstrumente. Ein Messinstrument, das im Sinne der klassischen Testtheorie eine möglichst hohe Reliabilität aufweist, ist möglicherweise für Verlaufsforschungen deshalb nicht geeignet, weil es zu „reliabel“ ist.

Die Verwendung von psychometrischen Verfahren, wie des Gießen-Testes, die explizit auf Veränderung der Wahrnehmung interpersonaler Bezüge hin konstruiert sind (BECKMANN u. RICHTER, 192??), ist für die psychoanalytische Verlaufsforschung deswegen nicht möglich, weil die wiederholte Anwendung des Messinstrumentes während der Behandlung eine Störvariable darstellen würde, über dessen Bedeutung als „Parameter“ im Sinne EISSLER's (1958) zwar klinisch sicher eine einhellige Meinung besteht – dass er untragbar wäre – während die Auswirkung solcher mit dem Behandlungsprozess intervenierender Maßnahmen noch kaum erforscht ist. Ohne Zweifel ist in den letzten Jahren hier eine langsam wachsende Risikobereitschaft vorhanden, die es z.B. ermöglicht, Patienten in analytischer Behandlung vierwöchentlich in Schlaflaboratorien elektroenzephalographisch zu untersuchen, ohne dass bisher für die Behandlung fatale Auswirkungen berichtet worden wären (s.d. KNAPP et al., 1975). Jedenfalls sind diese klinisch motivierten Bedenken ein wesentlicher Grund für unsere Überzeugung, dass Tonbandprotokolle für die Erforschung des Behandlungsprozesses deswegen so geeignet sind, weil ihre Herstellung – die Tonbandaufnahme selbst, nach allen bisher vorliegenden Berichten und Erfahrungen als Störvariable eine „quantité négligeable“ ist (s.d. WALLERSTEIN u. SAMPSON, 1971, S. 23).

Von den Kategorien des Wörterbuches wurden Veränderungen erwartet, die sich einerseits von nur zufälligen Schwankungen um einen festen Mittelwert herum unterscheiden, andererseits auch von den in Zeitserien erwarteten Trends zu differenzieren sind. Als Maß für die Variabilität der einzelnen Kategorien wurde zunächst der sogenannte Variationskoeffizient berechnet: $VK = \frac{s}{x}$.

Hierbei ergab sich als Folge der sehr unterschiedlichen Häufigkeiten der Kategorien der auffällige Befund, dass diejenigen Kategorien mit einem relativ kleinen Mittelwert einen ungewöhnlich hohen Variationskoeffizienten aufweisen. Diese Beziehung ließ sich korrelationsstatistisch sichern (Rang-Korrelation nach SPEARMAN ergibt einen $r_s = + 0.81$, sig. 1% Niveau).

So hat z.B. die Kategorie FOOD (7) einen VK von 11,67 und die Kategorie CLOTHING (8) einen VK von 4,74 beim Patienten, die Werte für den Analytiker sind ähnlich extrem.

Extrem hohe Variationskoeffizienten geben bei der vorliegenden Datenstruktur einen Hinweis darauf, dass der Mittelwert der Verteilung nicht wirklich ein mittleres Maß darstellt, um den

^{*)} Wir sind an anderer Stelle bereits kurz auf die Problematik solcher Zeitserien und ihrer Untersuchung eingegangen (SCHAUMBURG, KÄCHELE u. THOMÄ, 1974, S. 370); für eine ausführlichere Diskussion zur Problematik der Veränderungsmessung s. HARRIS, 1967).

herum sich die Werte gruppieren. Zur Überprüfung der Verteilung der einzelnen Kategorien wurde deshalb ein Häufigkeitshistogramm über die Stichprobe von 130 Stunden hergestellt^{*)}, auf der die Verteilung der Werte direkt graphisch abzulesen ist. Als Beispiel für eine recht typische Verteilungsform, die wir bei der Mehrzahl der Kategorien fanden, geben wir das Histogramm für die Kategorie FEMALE ROLE des Patienten wieder (Abb. 26)

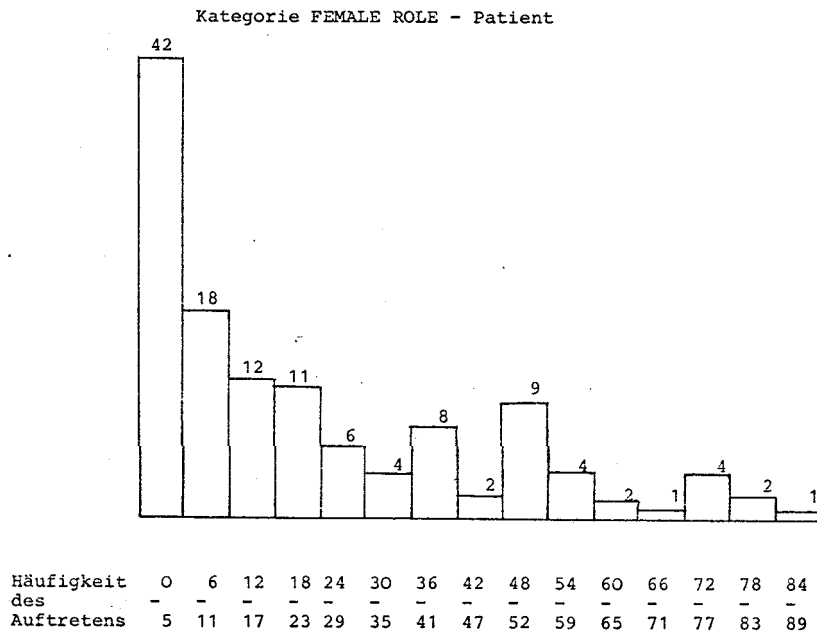


Abbildung 26

Aus der Abbildung 26 ist ersichtlich, dass die Kategorie FEMALE ROLE des Patienten in 42 der 130 Std. 0 – 5 mal auftritt; die übrigen 88 Stunden weisen einen Mittelwert von 35,4 Promille mit einer Standardabweichung von $\pm 33,98$ Promille auf. Die Ergebnisse der Verteilungsstudie legten es nahe, zunächst das Auftreten bzw. Nichtauftreten einer Kategorie zu bestimmen und die Berechnung des Variationskoeffizienten sinnvollerweise nur an jenen Stunden durchzuführen, in denen die Kategorie tatsächlich aufgetreten ist. Nur so konnten vergleichbare Hinweise auf die Variabilität gewonnen werden. Zu diesem Zwecke wurde zunächst die dichotome Verteilung der 45 Kategorien 1. Ordnung untersucht. Es zeigten sich recht aufschlussreiche Unterschiede in der Sprachverwendung zwischen Patient und Analytiker, die im folgenden beschrieben werden. Eine Übersicht der dieser Auswertung zugrundeliegenden dichotomen Häufigkeiten der 45 Kategorien 1. Ordnung gibt die Abbildung 27.

Besonders häufig vorkommende Kategorien sind die Kategorien DISTRESS, BODY PART, TIME REFERENCE, SPACE REFERENCE, NONSPECIFIC OBJECT, MESSAGE FORM, THOUGHT FORM, THINK; sie kommen beim Patient wie auch beim Analytiker in mehr als 75% aller Stunden vor. Eine größere Zahl von Kategorien kommt in ca. 50% der Stunden vor; so z.B. MALE ROLE, FEMALE ROLE, IDEAL WORLD u.a.

Einige Kategorien hingegen kommen in weniger als 25% der Stunden vor und sind hierdurch weniger als Verlaufsindikatoren geeignet, z.B. als Extrembeispiele die Kategorien FOOD und CLOTHING, dann aber auch DEFENSE MECHANISM, BAD, GOOD, EXPELL, GET, FOLLOW, AVOID, MOVE.

^{*)} Durchgeführt mit dem Programm BMD05 D, DIXON 1973.

Dichotome Häufigkeitsverteilung der Kategorien 1. Ordnung über 130 Stunden

Var.	Kategorien	Patient	Analytiker	chi ² df = 1	Signifikanz
1	Male Role	75	69		
2	Female Role	93	78		
3	Neuter Role	80	59	7,0	1 %
4	Job Role	43	21	12,5	1 %
5	Large Group	68	2	136,1	1 %
6	Body Part	108	85	7,6	1 %
7	Food	1	1		
8	Clothing	8	7		
9	Tool	32	37		
10	Natural Object	42	26	8,6	1 %
11	Non Spec. Object	12	117		
12	Time Ref.	129	117		
13	Space Ref.	103	114		
14	Quantity Ref.	81	64		
15	Social Place	94	82		
16	Natural World	48	52		
17	Ideal World	60	79	4,7	5 %
18	Deviation	73	14	79,3	1 %
19	Action Norm	47	46		
20	Message Form	111	119		
21	Thought Form	128	128		
22	Arousal	93	97		
23	Urge	65	69		
24	Affection	88	30	55,5	1 %
25	Pleasure	83	65	10,8	1 %
26	Distress	129	120		
27	Anger	79	68		
28	Sense	70	85		
29	Think	111	109		
30	If	47	15	29,1	1 %
31	Not	34	25		
32	Defense Mechanism	18	16		
33	Good	13	34	18,7	1 %
34	Bad	4	10		
35	Communicate	53	57		
36	Approach	33	38		
37	Guide	35	27		
38	Follow	15	16		
39	Attack	55	72	5,1	5 %
40	Avoid	10	28	17,1	1 %
41	Attempt	56	45		
42	Get	21	26		
43	Expel	13	35	18,7	1 %
44	Work	47	37		
45	Move	25	34		

Abbildung 27

Folgende Kategorien 1. Ordnung treten in den 130 Stunden signifikant verschieden bei Patient und Analytiker auf:

Der Patient beschäftigt sich häufiger mit den Rollenkategorien NEUTER ROLE und JOB ROLE; er adressiert damit häufiger nicht geschlechtsspezifisch zugeordnete Personen sowie die Person des Arztes bzw. der Krankenschwester. Praktisch nie verwendet der Analytiker das Wort „Leute“, was beim Patient in 52% der Stunden vorkommt. Mit „Leuten“ apostrophiert der Patient oft nicht näher bezeichnete Personen, an denen sich seine Angst bereits gedanklich festmacht. Praktisch und theoretisch wichtig ist auch die signifikante Zurückhaltung des Analytikers bezüglich der Kategorie DEVIATION; d.h. der Analytiker spricht nur in 10 % der Stunden direkt über körperbezogene Krankheitserscheinungen (wie z.B. Tachykardie, Anfall, Pulsjagen), die beim Patient in 55% der Stunden auftreten. Dies erscheint uns ein Hinweis auf die selektive Aufmerksamkeit bzw. selektive Vermeidung in der Wortwahl des Analytikers. Fast dreimal so oft wie der Analytiker spricht der Patient über Nähe, Zärtlichkeit und Sehnsucht (AFFECTION). Die Kategorie PLEASURE, mit der die Lust/Unlustäußerungen des Patienten erfasst werden, tritt ebenfalls signifikant häufiger beim Patient auf. Der Analytiker hingegen akzentuiert die Kategorien IDEAL WORLD und GOOD, womit er offensichtlich einen Wertebereich anspricht. Weiterhin findet sich eine signifikante Verstärkung seines sprachlichen Handelns im Vergleich zum Patient bei drei Kategorien (ATTACK, AVOID und EXPELL), mit denen der Analytiker aggressives Handeln thematisiert. Auch hier finden sich deutliche Hinweise auf das theorie-bezogene sprachliche Handeln des Therapeuten.

Bei der Bewertung der hier aufgeführten Unterschiede muss berücksichtigt werden, dass hier nur die dichotomisierten Häufigkeiten ausgewertet wurden.

Im Anschluss an die Dichotomisierung der Kategorienhäufigkeiten wurden für diejenigen Stunden erneut der Mittelwert, die Standardabweichung und der Variationskoeffizient bestimmt, in denen die jeweilige Kategorie den Wert größer als Null erhalten hatte.

Wir führten damit implizit eine Unterscheidung im Messniveau der inhaltsanalytischen Daten ein. Nach unseren bisherigen Erfahrungen mit solchen Daten könnte es sinnvoll sein, zunächst das Auftreten bzw. Nicht-Auftreten einer Kategorie und erst im Anschluss daran die quantitative Ausprägung zu bestimmen. Damit wäre eine Gewichtung der Art vorgenommen, dass die ja/nein-Entscheidung von größerer Bedeutung als die quantitative Ausprägung einer Kategorie ist.

Das Ergebnis der erneuten Berechnung des Variationskoeffizienten ergab, dass der Variationskoeffizient für die meisten Variablen nun Werte zwischen 0.91 und 0.61 annahm. Am Beispiel einiger Kategorien verdeutlichen wir im folgenden das Ergebnis (Abbildung 28).

Kategorie		Variations- koeffizient - Patient -	N Std.	Variations- koeffizient - Analytiker -	N Std.
1	MALE ROLE	0,85	75	0,86	69
2	FEMALE ROLE	0,95	93	0,94	78
3	NEUTER ROLE	0,80	80	0,75	59
25	PLEASURE	0,99	83	1,31	65
26	DISTRESS	0,61	129	0,59	120
27	ANGER	0,96	79	0,98	68

Abbildung 28

Diese Kategorien – stellvertretend für die anderen Kategorien des Wörterbuchs – weisen ausgeprägte Variationskoeffizienten auf. Das Ergebnis lässt sich so interpretieren, dass in den Stunden, in denen eine Kategorie auftritt eine recht breite Variabilität der Häufigkeitswerte auftritt. Ausnahmen hiervon gibt es nur wenige. Beim Patient weisen z.B. die Kategorien 11 (NON SPECIFIC OBJECT), 12 (TIME REFERENCE) und 13 (SPACE REFERENCE) einen VK von kleiner als 0,60 auf, beim Analytiker ist die Kategorie 4 (JOB ROLE) deutlich weniger variiert.

Zweck der Berechnung des Variationskoeffizienten war es, diejenigen Kategorien bestimmen zu können, die sich besonders gut als Indikatoren für Veränderungen des Sprachinhaltes eignen würden. Als zusammenfassendes Ergebnis müssen wir jedoch festhalten, dass fast alle Kategorien – bis auf wenige Ausnahmen – bedeutsame Variationskoeffizienten aufweisen, die die Brauchbarkeit der Kategorien als Indikatoren für Veränderung unterstreichen. Hierzu war es allerdings notwendig, die Verteilung der Kategorien nur in jenen Stunden zu untersuchen, in denen sie auch faktisch auftreten. Die erhebliche Unterschiedlichkeit der dichotomen Ja/Nein-Verteilung der einzelnen Kategorien lässt es ratsam erscheinen, bei speziellen späteren Untersuchungen diejenigen Kategorien zu eliminieren, die nur in wenigen der insgesamt 130 Stunden überhaupt auftreten.

4.2.4 Die Ähnlichkeit des Sprachinhaltes bei Patient und Analytiker

Bei der vergleichenden Betrachtung der Mittelwerte für die 61 Kategorien von Patient und Analytiker hatten wir eine auffallende Ähnlichkeit der Mittelwerte festgestellt. Nun lässt die Ähnlichkeit der mittleren Häufigkeitswerte keine Schlüsse darüber zu, ob die Verwendung der Kategorien innerhalb bestimmter Zeiträume bei Patient und Analytiker ähnlich ist. Wie bereits oben erwähnt, beschrieb LAFFAL (1967), dass sich der Sprachinhalt zweier Sprecher mehr ähnelt, als der Sprachinhalt des gleichen Sprechers in verschiedenen Gesprächen. Die Zentrierung eines Gespräches auf ein Thema erzwingt gewissermaßen eine Angleichung des Sprachinhaltes. LAFFAL verwendete jedoch für die Bestimmung der Ähnlichkeit ein sehr globales Maß, indem er die Kategorienstruktur der beiden Sprecher insgesamt vergleicht. Im folgenden stellen wir die Untersuchung der Ähnlichkeit der beiden Gesprächspartner in diesem therapeutischen Dialog anhand der Ähnlichkeiten des Verlaufs der einzelnen Kategorien in den Mittelpunkt. Dabei lassen sich besser die Gemeinsamkeiten und Unterschiede herausarbeiten als dies bei einem globalen Vergleich möglich ist.

Die bei der Untersuchung der dichotomen Häufigkeitsverteilung festgestellten Ähnlichkeiten bzw. signifikanten Unterschiede können die hier gestellte Frage nicht beantworten. Es ist ja denkbar, dass, immer wenn der Patient eine bestimmte Kategorie A verwendet, der Analytiker aus behandlungstechnischen Gründen diese nicht verwendet und umgekehrt. Um die Ähnlichkeit des Sprachinhaltes über den Behandlungsverlauf hinweg zu prüfen, wurden deshalb die Kategorien über die 130 Stunden miteinander korreliert.^{*)}

Für die Fragestellung wurden die zwei Datensätze von Patient und Analytiker über die 130 Stunden zusammengefügt und einer gemeinsamen Korrelationsberechnung unterzogen.^{*)}

Die folgende Tabelle gibt die Korrelation zwischen den Kategorien von Patient und Analytiker für die beiden Stichproben an (Abb. 29).

^{*)} Gerechnet mit BMD02 D; (DIXON, 1973).

^{*)} Gerechnet mit BMD02 D; (DIXON, 1973).

Korrelation ausgewählter Variablen zwischen Patient und Analytiker über die Stichprobe I (130 Std.)

Variable		S I
1	MALE ROLE	0.64
2	FEMALE ROLE	0.59
3	NEUTER ROLE	0.39
4	JOB ROLE	0.39
6	BODY PART	0.48
9	TOOL	0.20
10	NATURAL OBJECT	0.45
11	NON SPECIFIC OBJECT	0.06
12	TIME REF.	0.34
15	SOCIAL PLACE	0.28
16	NATURAL WORLD	0.23
17	IDEAL WORLD	0.38
18	DEVIATION	0.44
19	ACTION FORM	0.15
20	MESSAGE FORM	0.35
21	THOUGHT FORM	0.30
22	AROUSAL	0.28
23	URGE	0.31
24	AFFECTION	0.25
25	PLEASURE	0.75
26	DISTRESS	0.33
27	ANGER	0.66
28	SENSE	0.12
29	THINK	0.24
30	IF	0.22
31	NOT	0.07
35	COMMUNICATE	0.25
36	APPROACH	0.42
37	GUIDE	0.26
38	FOLLOW	0.28
40	AVOID	0.35
41	ATTEMPT	0.40

$p * 0.01, r] 0.22$

$p * 0.001, r] 0.28$

Abbildung 29

Im Überblick fällt auf, dass die Mehrzahl der Kategorien eine statistisch signifikante Ähnlichkeit des Verlaufes über die 130 Stunden aufweist. Allerdings ist diese Signifikanz inhaltlich nur wenig aussagekräftig, denn sie wird zum größten Teil durch das hohe N der untersuchten Stichprobe erzeugt.

Von einer auffallenden Ähnlichkeit des Kategorienverlaufes kann man jedoch bei der Kategorie PLEASURE sprechen, die mit + 0.75 bei Patient und Analytiker sehr eng miteinander

kovariiert. Da es sich hierbei um eine Kategorie handelt, die bei Patient und Analytiker auch einen positiven Zeittrend aufweist^{*)} (d.h. sie nimmt im Verlauf der Behandlung zu), wird mit dieser Kategorie vermutlich ein Behandlungsinhalt angesprochen, der sich als langfristige systematische Veränderung ansprechen lassen wird. Es ist in diesem Zusammenhang von besonderem theoretischem Interesse, dass auch die Kategorie ANGER sehr deutlich positiv kovariiert (+ 0.66), wobei sich zusätzlich aus der gemeinsamen Korrelationsmatrix der 2 x 61 Kategorien entnehmen lässt, dass auch die PLEASURE- und ANGER-Kategorien positiv miteinander korrelieren.

Im Gegensatz zu diesen Kategorien weist die Angst-Kategorie DISTRESS nur eine geringe Kovariation auf; die Korrelation erklärt nur 11% der gemeinsamen Varianz der Kategorie bei Analytiker und Patient auf. Für die mit dieser Kategorie erfasste Verbalisierung von Angst, Ekel, Beschwerden und Schwierigkeiten scheint sich kein Abstimmungsprozess hergestellt zu haben; dies weist darauf hin, dass bei den Klagen des Patienten der Analytiker nicht die Klagen unverändert aufgreift, sondern offensichtlich Neuformulierungen angeboten hat. Hier sind noch weitere detaillierte Untersuchungen des Austauschprozesses notwendig, um die bedeutungstechnische Bedeutung der Ähnlichkeit bzw. Verschiedenheit des Sprachinhaltes voll auszuschöpfen.

Mit diesem Zwischenergebnis sind einige Beschreibungen der mit Hilfe des EVA-Programmsystems gewonnen Rohdaten abgeschlossen. Es wurde gezeigt, dass die inhaltsanalytischen Werte für die Stichprobe von 130 Stunden einen differenzierten und variablen Datensatz ergeben, mit dem Beschreibungen des Behandlungsprozesses durchgeführt werden können. Von den vielfältigen Ansätzen, die zum gegenwärtigen Zeitpunkt den exploratorischen Charakter der Anwendung der maschinellen Inhaltsanalyse kennzeichnen, werden im folgenden Kapitel auf die Vorhersage klinischer Konzepte durch inhaltsanalytische Variablen ausführlich eingegangen. Andere Ansätze zur Operationalisierung klinischer Angstkonzepte wurden bereits anderswo beschrieben (KÄCHELE, GRÜNZIG, BÜSCHER, 1976; KÄCHELE, GRÜNZIG, MERGENTHALER, 1976) bzw. werden zur Zeit verfolgt. Hierbei experimentieren wir mit faktorenanalytischen Beschreibungsmodellen, um die Informationsvielfalt auf wenige empirisch gewonnene Dimensionen zu reduzieren. Diskriminanzanalytische Verfahren werden eingesetzt, um für verschiedene Behandlungsphasen typische Kategorienkonstellationen zu bestimmen.

^{*)} Die Zeittrends der Kategorien wurden mit einem speziellen Programm TRENDEL berechnet, welches anschließend auch eine Trendelimination durchführt. Das Programm wurde von Dr. Grünzig geschrieben.

4.3 Die Vorhersage klinischer Konzepte durch die maschinelle Inhaltsanalyse

4.3.1 Das Verfahren der multiplen Regression als Vorhersagemodell

4.3.2 Charakterisierung der klinischen Konzepte

4.3.3 Die Ergebnisse der Regressionsanalyse

4.3.4 Zur klinischen Prüfung der Ergebnisse

4.3.5 Die Prüfung der Generalisierbarkeit der Ergebnisse

4.3.6 Der Verlauf maschinell ermittelter klinischer Konzepte

4.3.7 Kritische Würdigung der Ergebnisse

4.3 Die Vorhersage klinischer Konzepte durch die maschinelle Inhaltsanalyse

Als Grund für die Entwicklung von Prozeduren zur maschinellen Inhaltsanalyse haben wir eingangs die forschungsökonomisch sehr ungünstige Situation beschrieben, dass mit der Verwendung von Verbatim-Protokollen eine manuell kaum zu bewältigende Fülle von verbalen Daten anfällt. Die auf solche Daten angewiesene psychoanalytische Prozessforschung kann sich nur weiterentwickeln, wenn eine Computer- und Softwaretechnologie zur Verfügung steht, die die Auswertung unterstützt.

Dies trifft für die Sozialwissenschaften insgesamt zu, bei der mit der Entwicklung der Datenverarbeitungsanlagen bis dahin kaum bearbeitbare Fragen aufgegriffen werden konnten. Die Erstellung zentraler Datenarchive für die Ergebnisse von multinationalen Umfragen ist hierfür wohl das bekannteste Beispiel (in Deutschland exemplarisch durch das Zentralarchiv für empirische Sozialforschung in Köln vertreten).

Die Aufgabe der maschinellen Inhaltsanalyse kann es nicht sein, alle Schritte sozialwissenschaftlichen Forschungsprozesses zu übernehmen, aber es sollte möglich sein, einige Arbeitsschritte den Datenverarbeitungsanlagen zu übertragen, für die der „stupid clerk“ weit aus geeigneter ist als die menschliche Intelligenz.

„Eine wesentliche Aufgabe jeder empirischen Wissenschaft ist es zweifellos, die ungeheure Fülle vorliegender empirischer Informationen zu reduzieren. Da unsere menschliche Intelligenz schon die Fülle und die Komplexität der Daten einer einzigen größeren Untersuchung nicht adäquat zu reduzieren vermag – so würden wir nach dem Lesen der Interviews (standardisiert oder nicht standardisiert) von nur 100 Personen wohl kaum eine gezielte Frage nach dem Ergebnis dieser Untersuchung sinnvoll beantworten können –, ist es notwendig, bestimmte Hilfsmittel zu verwenden“ (KRIZ, 1975, S. 130).

Im Mittelpunkt der klinischen Prozessforschung steht die Untersuchung von Veränderungen solcher grundlegender Phänomene oder Syndrome, die wegen ihrer klinischen Bedeutung konzeptualisiert werden – hier im folgenden kurz ‚klinische Konzepte‘ genannt. Wir sehen es als ein Ziel dieser Untersuchungen, solche Veränderungen klinischer Konzepte durch die Prozedur der maschinellen Inhaltsanalyse zu erfassen.

Um dies zu ermöglichen, ist es notwendig, die klinischen Konzepte so zu operationalisieren, dass eine Arbeitsanweisung in Form eines Programmsystems Schritt für Schritt die klinische Beurteilung nachvollziehen kann. Noch gibt es aber wenig detaillierte Kenntnis darüber, welche mikropsychologischen und linguistischen Vorgänge bei der Beurteilung eines klinischen Konzepte wie z.B. Kastrationsangst im Beurteiler ablaufen. Eine vollständige Theorie darüber, aus der deduktiv Programmschritte ableitbar wären, liegt noch nicht vor; über erste Schritte, einige der klinischen ‚cues‘ zu entdecken, sind wir noch nicht hinaus (SPENCE u. LUGO, 1973; DAHL, 1972, 1974).

Die Gründe hierfür liegen nicht nur in der kaum überschaubaren Komplexität der klinischen Urteilsbildung. Sie sind auch darin zu suchen, dass es bisher kaum empirische Ansätze gegeben hat, das klinisch-psychoanalytische Denken und die Vorgänge bei der klinischen Urteilsbildung systematisch zu untersuchen (s.d. COHEN, 1969).

Eigene Untersuchungen zur Wahrnehmungs- und Urteilsstrategie des Klinikern anhand von exzerpierten Äußerungen von Patienten und Analytikern haben gezeigt, dass es über vereinfachende Annahmen möglich ist, wesentliche verbale Elemente der Äußerungen zu identifizieren, die die Kliniker vermutlich zu ihrer Beurteilung veranlasst haben (BÖCKENFÖRDE, GRÜNZIG, KÄCHELE u. THOMÄ, 1976; GRÜNZIG u. KÄCHELE, 1976; KÄCHELE, GRÜNZIG und MERGENTHALER, 1976).

Im folgenden wird ein Ansatz beschrieben, der sich zur Operationalisierung klinischer Konzepte eines speziellen statistischen Modells bedient.

4.3.1 Das Verfahren der multiplen Regression als Vorhersagemodell

Um ein Phänomen aus der Kenntnis anderer Phänomene X_1, X_2, \dots, X_n vorherzusagen, ist es notwendig, regelhafte, aber nicht notwendigerweise kausale Beziehungen der beiden Phänomene Y und X zu kennen^{*)}. Da wir von der Annahme ausgehen, dass sich die klinischen Urteile auf sprachliche Inhalte stützen, die aus den Verbatim-Protokollen zu entnehmen sind, müsste es möglich sein, jene verbalen Indikatoren für klinische Konzepte zu bestimmen, auf die sich der klinische Beurteiler im wesentlichen stützt.

Hierzu bietet sich das Modell der multiplen linearen Regression an. In dieser Modellvorstellung wird angesetzt, dass die abhängige Variable Y aus folgender Linearkombination der unabhängigen Variablen X_i resultiert:

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n + C$$

Die Werte a_i stellen Gewichtungsfaktoren dar, mit denen die jeweiligen X_i multipliziert werden, um zu einer besten Vorhersage von Y zu gelangen; C stellt eine additive Konstante dar. Bei gegebenen Y und X_i sind die Koeffizienten a_i so zu wählen, dass die ‚Nichterklärbarkeit‘ von Y aus X_i ein Minimum wird, d.h. dass die aus X_i erklärbare Varianz von Y ein Maximum wird. Dieses Problem ist mit den mathematischen Mitteln der Differentialrechnung eindeutig lösbar (GAENSSLEN u. SCHUBÖ, 1973).

Wenn wir die klinischen Konzepte, deren Verläufe wir an einer größeren Stichprobe von 130 Stunden untersuchen wollen, als abhängige Variable Y ansetzen, so muss eine Linearkombination ausgewählter Inhalts-Kategorien gefunden werden, die eine optimale Vorhersage der abhängigen klinischen Variablen ergibt.

Zur Durchführung dieser Operationalisierungsprozedur werden aber zur Ermittlung der prädiktiv potenten Inhaltskategorien zunächst ein Satz klinisch beurteilter Stunden benötigt. Im Rahmen der klinischen Verlaufsuntersuchung des hier analysierten Behandlungsfalles wurden 55 Stunden mit Schätzskalen hinsichtlich verschiedener klinischer Konzepte beurteilt; diese Stunden haben wir als Ausgangsstichprobe verwendet. Um die unübersichtliche Vielfalt der Inhaltskategorien auf die für diese Vorhersage relevanten Kategorien zu reduzieren, wurden schrittweise multiple Regressionsanalysen gerechnet (mit BMD02R, DIXON, 1973), und zwar wurde jeweils diejenige Inhaltskategorie in die Regressionsgleichung neu hinzugenommen, die unter den noch nicht berücksichtigten Kategorien den höchsten Anteil an der Vorhersage leistet. Als Kriterium für den Abbruch dieses iterativen Vorgehens wurde die Nicht-Signifikanz des Regressionskoeffizienten derjenigen Inhaltskategorie herangezogen, die im folgenden Schritt in die Regressionsgleichung hätte aufgenommen werden sollen.

Die Ergebnisse dieses Vorgehens können als rein beschreibende Statistik für die Beziehung zwischen den Inhaltskategorien in den untersuchten Stunden und der Kriteriumsvariable, nämlich dem klinischen Konzept, angesehen werden.

Betrachtet man die untersuchten Stunden aber als Zufallsstichprobe aus der Gesamtpopulation aller Behandlungsstunden, so muss der multiple Korrelationskoeffizient R einer zufalls-kritischen Prüfung unterzogen werden, um mit einer definierten Irrtumswahrscheinlichkeit von der Stichprobe auf die Population übertragbar zu sein.

^{*)} HEMPEL und OPPENHEIM (1953) haben die Beziehung zwischen Vorhersage und Erklärung in dem nach ihnen benannten Schema systematisiert. Eine ausführliche Diskussion der strukturellen Identität von Vorhersage und Erklärung findet sich bei STEGMÜLLER, 1969, S. 153-199.

4.3.2 Charakterisierung der klinischen Konzepte

Zum besseren Verständnis der multiplen Regressionsanalyse und ihrer Ergebnisse beschreiben wir kurz die klinisch-systematische Untersuchung des Behandlungsverlaufes durch die Verwendung von Beurteilungsskalen. Die ausführliche Darstellung dieses Vorgehens wurde bereits an anderer Stelle gegeben^{*)}. Im Mittelpunkt der seelischen Schwierigkeiten des hier untersuchten Patienten steht das Erleben der Angst in ihren vielfältigen Formen. Von daher lag es nahe, bei der klinischen Verlaufsbeschreibung und skalierten Beurteilung psychoanalytische Angstkonzepte als organisierende und selegierende Strukturen zu verwenden. Für die Einstufungsuntersuchung wurden in Anlehnung an GOTTSCHALK u. GLESER (1969) sechs Angstkonzepte verwandt; es wurde jeweils eine Schätzskala verwandt, auf der die Intensität der Angstthematik, d.h. ihre psychodynamische Wirksamkeit in der Behandlungssituation, eingeschätzt werden sollte. Da klinisch zwischen bewussten und unbewussten Gefahrensituationen unterschieden wird, wurde eine weitere Skala „Grad der Bewusstheit“ eingeführt, auf der im Sinne des topographischen Gesichtspunktes über den Ort der Angst jeweils entschieden werden musste.

Neben den Angstkonzepten wurde eine Beurteilung der Übertragungskonstellation durchgeführt, wobei wir aus Gründen der Handbarkeit nur zwischen der ‚freundlichen‘ und ‚feindseligen‘, der libidinösen und aggressiven Übertragung unterschieden haben.

Es wurden außerdem vier Konzepte zur Beurteilung der Arbeitsbeziehung einbezogen, auf die wir hier nicht eingehen, da sie im weiteren Verlauf dieser Untersuchung nicht verwendet werden.

Aus verschiedenen Gründen, die wir bereits weiter oben (Kap. 3) bei den Problemen der Stichprobenziehung diskutiert haben, wurde für diese klinische Einstufungsuntersuchung eine Stichprobe ausgewählt, die aus den Stunden 1 - 5, 51 - 55,, 501 - 505 besteht. Diese 55 Stunden wurden von drei Beurteilern^{*1)} anhand der transkribierten Verbatim-Protokolle hinsichtlich der sechs Angstkonzepte und der beiden Übertragungskonzepte beurteilt. Für die Bestimmung der Inter-Beurteiler-Übereinstimmung wurden die Intensitäts- und Bewusstheitsskalen durch einen einfachen Abbildungsformalismus^{*2)} auf einer einzigen, gemeinsamen Skala abgebildet. Hierbei wurde der Überlegung Rechnung getragen, dass die Einschätzung der dynamischen Wirksamkeit eines Konzeptes der Beurteilung der Bewusstheit logisch und psychologisch vorgeordnet ist.

Die folgende Tabelle (Abb. 30, S. 94) enthält das Ausmaß der Inter-Beurteiler-Übereinstimmungen in Form von Produkt-Moment-Korrelationskoeffizienten zwischen den jeweiligen Beurteilerpaaren wie auch für die Beurteilergruppe insgesamt. (Das Todesangstkonzept bleibt in der weiteren Darstellung unberücksichtigt, da dieses Konzept übereinstimmend nur selten als dynamisch relevant angesehen wurde).

^{*)} s.d. DFG-Forschungsbericht 1973, Teil B: Klinische Verlaufsbeschreibung sowie die überarbeitete Fassung in KÄCHELE et al. (1976).

^{*1)} Die Beurteiler waren aus praktischen Gründen mit der damaligen Forschungsgruppe identisch: A = der behandelnde Psychoanalytiker, B = ein Arzt in psychoanalytischer Ausbildung und C = eine klinische Psychologin.

^{*2)} $10 \cdot 3 \cdot X_{\text{Int.}} + 1 \cdot X_{\text{Bew.}} = X_{\text{g}}$

Tabelle:Inter-Beurteiler Übereinstimmung
(Produkt-Moment- Korrelation)

	A/B	A/C	B/C	r ²
Kastrations-Angst	.31	.38	.29	.33
Trennungs-Angst	.55	.53	.30	.46
Schuld-Angst	.37	.34	.31	.34
Beschämungs-Angst	.29	.04	.07	.13
diff. Angst	.59	.53	.49	.54
lib. Übertragung	.44	.59	.32	.45
aggress. Übertragung	.47	.35	.42	.41

N = 55

p * .05 bei r] .27

p * .01 bei r] .35

Abbildung 30

Die Homogenitätsprüfung der Koeffizienten der drei Beurteilerpaare ergab in keinem Fall eine bedeutsame Abweichung.

Die Übereinstimmung zwischen den drei Beurteilern war – würdigt man die Aufgabe und die Komplexität des Beurteilungsvorganges – recht zufriedenstellend und liegt durchaus im oberen Bereich, der bisher von anderen Autoren (BELLAK u. SMITH, 1956; STRUPP, CHASSAN und EWING 1966)^{*)} für solche klinischen Konzepte angegebenen Höhe.

Führt man sich aber die Aussagekraft solcher Korrelationskoeffizienten – selbst bei zufallskritischer Absicherung – deutlicher vor Augen, so bedeutet die maximale Übereinstimmung in Höhe von $r = +0.59$ (Konzept „Diffuse Angst“ zwischen Beurteiler A und B), dass diese Beurteiler nur zu etwa 36 % ein gemeinsames Beurteilverhalten aufweisen. Mit anderen Worten, es ist der Grad der Nicht-Übereinstimmung immer noch höher als der Grad der Übereinstimmung.

Trotzdem bleibt, wie eben skizziert, der Grad der Übereinstimmung problematisch. Deshalb haben wir im Anschluss an die Berechnung der Inter-Rater-Übereinstimmung die faktorielle Struktur der Beurteilungen untersucht. Die Methode der Faktorenanalyse soll als deskriptives Verfahren die gemeinsame Varianz der Variablen auf Strukturen höherer Ordnung als die der Variablen zurückführen:

„Bei der eigentlichen Faktorenanalyse sucht man die beobachteten Variablen auf Faktoren (F_i) zurückzuführen, die allen Variablen gemeinsam sind; daneben werden Faktoren definiert, die nur für eine einzige Variable von Bedeutung sind“ (BAUMANN, 1974, S. 84).

Ohne hier auf einschränkende Überlegungen eingehen zu können, die sich auf die Anwendung der Faktorenanalyse auf solche Daten beziehen, geben wir im folgenden Auszüge aus den Ergebnissen einer Faktorenanalyse, die nach der Hauptachsen-Methode mit nachfolgender Varimax Rotation gerechnet wurde (mit BMD08M, DIXON, 1973).
(Abbildung 31, S. 95)

^{*)} Eine ausführliche Diskussion der hier anstehenden Probleme und eine Literaturübersicht findet sich bei BÖCKENFÖRDE, GRÜNZIG, KÄCHELE u. THOMÄ, 1976).

Auszug aus der rotierten FaktorenmatrixFaktor 2Trennungsthematik bei libidinöser Übertragung

A	Lib. Übertragung	+ 0.67
A	Trennungsangst	+ 0.79
B	Lib. Übertragung	+ 0.43
B	Trennungsangst	+ 0.59
C	Lib. Übertragung	+ 0.79
C	Trennungsangst	+ 0.66

Faktor 3Diffuse Angstthematik bei aggressiver Übertragung

A	Aggress. Übertragung	+ 0.67
A	Diffuse Angst	+ 0.74
B	Aggress. Übertragung	+ 0.61
B	Diffuse Angst	+ 0.64
C	Aggress. Übertragung	+ 0.35
C	Diffuse Angst	+ 0.63

Faktor (spezifischer Faktor einer neun-faktorellen Lösung)Kastrationsangst

A	Kastrationsangst	+ 0.70
B	Kastrationsangst	+ 0.68
C	Kastrationsangst	+ 0.72

Abbildung 31

Aus der Fünf-Faktoren-Lösung haben wir hier zwei Faktoren dargestellt, die die vereinheitlichende Funktion des faktoren-analytischen Ansatzes deutlich hervorheben. Der Faktor 2 beschreibt eine deutliche Kovarianz der Beurteilung von libidinöser Übertragung als freundlicher Beziehungseinstellung zum Analytiker im Verbund mit der Thematisierung der Trennungsproblematik.

Der Faktor 3 vereinigt ebenfalls eine Angstthematik mit einer der Übertragungsbeurteilung. Es ist offensichtlich, dass die Beurteiler die auf der objektlosen Stufe der Angst kreisenden Äußerungen des Patienten, die sich in manchen Stunden fast salvenartig wiederholen, im

Kontext einer aggressiven = feindselig-beschimpfenden Einstellung zum Analytiker beurteilt haben.

Bei der Fünf-Faktoren-Lösung, die wir endgültig gewählt haben, weisen die drei Kastrationsangst-Variablen der Beurteiler auf keinem der fünf Faktoren eine signifikante Ladung auf. Bei einer versuchsweise gerechneten Neun-Faktoren-Lösung erscheinen die drei Variablen – wie zu erwarten^{*)} – als eigenständiger spezifischer Faktor.

Wir haben uns in der Darstellung der Ergebnisse der faktorenanalytischen Untersuchung der Einstufungen auf die drei klinischen Konzepte bereits beschränkt, die bei der hier zu beschreibenden Vorhersage-Untersuchung verwendet wurden. Wir haben diese drei deshalb ausgewählt, weil sie unserer Erachtens gut identifizierbare sprachliche Referenzen aufweisen mussten.

Für die Vorhersage-Studie wurden die Faktorenscores der drei Konzepte verwandt, die für die 55 untersuchten Stunden berechnet worden waren.

4.3.3 Die Ergebnisse der Regressionsanalyse

Die inhaltsanalytischen Datensätze von je 61 Kategorien für Patient und Analytiker für 55 Stunden wurden aus der Stichprobe I (N = 130 Stunden) entnommen. Die Werte für die Kategorien wurden in dichotomisierter Form verwendet, um die mit der schiefen Verteilung zusammenhängenden methodischen Probleme zu umgehen. Dabei wurde jeweils am Mittelwert einer Kategorie dichotomisiert, was für die Kategorien mit geringen Häufigkeitswerten praktisch einer Dichotomisierung zwischen Null und 1 gleichkommt, bei den Kategorien mit durchschnittlich höheren Auftretenshäufigkeiten jedoch eine gewisse Berücksichtigung des Mittelwertes der Kategorie erlaubt.

In verschiedenen Voruntersuchungen wurden zunächst für Patient und Analytiker getrennt diejenigen Variablen ermittelt, die für die Vorhersage der klinischen Variablen vermutlich relativ viel beitragen. Hierbei wurden insgesamt 47 inhaltsanalytische Kategorien ermittelt (24 P-Variablen und 23 A-Variablen), die dann endgültig mit den drei klinischen Konzepten einer schrittweisen multiplen Regression unterzogen wurden (Abb. 32).

^{*)} Aus der Korrelationsmatrix ist zu entnehmen, dass die drei Variablen deutlich miteinander positiv kovariieren.

Liste der inhaltsanalytischen Variablen für die endgültige multiple Regressionsanalyse

Patient	Mittelwert	Analytiker	Mittelwert
FEMALE ROLE	0.36	MALE ROLE	0.27
NATURAL WORLD	0.33	JOB ROLE	0.15
DEVIATION	0.38	BODY PART	0.38
MESSAGE FORM	0.35	TOOL	0.33
THOUGHT FORM	0.38	TIME	0.38
AROUSAL	0.44	SOCIAL PLACE	0.31
URGE	0.31	NATURAL WORLD	0.33
PLEASURE	0.33	DEVIATION	0.11
DISTRESS	0.51	THOUGHT FORM	0.45
ANGER	0.31	AROUSAL	0.42
SENSE	0.35	AFFECTION	0.23
NOT	0.33	URGE	0.38
GOOD	0.11	DISTRESS	0.43
COMMUNICATE	0.31	ANGER	0.27
GUIDE	0.25	THINK	0.42
ATTACK	0.38	IF	0.13
EXPEL	0.13	NOT	0.13
HIGH STATUS	0.33	DEFENSE	0.13
PEER STATUS	0.33	APPROACH	0.25
LOW STATUS	0.31	ATTACK	0.31
OVERSTATE	0.22	MOVE	0.29
SIGN STRONG	0.29	PEER STATUS	0.20
SIGN AFFECT	0.31	UNDERSTATE	0.24
SEX SENSE	0.31		

Abbildung 32

Für jede Kategorie wurde zusätzlich der Mittelwert angegeben. Da es sich um dichotomisierte Daten handelt, zeigt der Mittelwert den Prozentsatz der Stunden an, in denen die Kategorie aufgetreten ist. In der Mehrzahl treten die ausgewählten Kategorien in ca. 20 - 40% der 55 Stunden auf. Kategorie des Patienten wie ‚GOOD‘ wurden trotz des relativ seltenen Auftretens beibehalten, da sie in den Voruntersuchungen eine relativ hohe prädiktive Kraft für ein klinisches Konzept aufwiesen.

Im folgenden werden die Ergebnisse der schrittweisen multiplen Regressionsanalyse für die drei klinischen Konzepte dargestellt.

4.3.3.1 Vorhersage des klinischen Konzeptes 1 „Kastrationsangst“

Die schrittweise Regressionsanalyse ergab folgende inhaltsanalytische Kategorien von Patient und Analytiker, die eine gemeinsame multiple Korrelation mit dem Konzept „Kastrationsangst“ von + 0.92 aufweisen (Abb. 33):

Schritt Nr.	Variable		multuples R
1	THOUGHT FORM	- A	0.36
2	ANGER	- P	0.46
3	AFFECTION	- A	0.53
4	MESSAGE FORM	- P	0.58
5	DEVIATION	- P	0.64
6	GUIDE	- P	0.69
7	SEX THEME	- P	0.75
8	FEMALE ROLE	- P	0.78
9	SIGN AFFECT	- P	0.81
10	COMMUNICATION	- P	0.83
11	DISTRESS	- A	0.85
12	THINK	- A	0.87
13	DEFENSE	- A	0.90
14	MOVE	- A	0.91
15	UNDERSTATE	- A	0.92

Abbildung 33

Mit diesen 15 Kategorien von Patient und Analytiker lassen sich 84,6% der Varianz des klinischen Konzeptes „Kastrationsangst“ aufklären. Zur Bestimmung der vorhergesagten Werte gehen die 15 Kategorien allerdings mit unterschiedlicher Gewichtung in die Regressionsgleichung ein. Zur inhaltlichen Überprüfung dieser Variablen geben wir nun die jeweiligen Regressionskoeffizienten der Variablen an, die über die Gewichtung der Kategorien Auskunft geben (Abb. 34, S. 98).

Patientenvariable	Regressions-Koeffizient	Analytikervariable	Regressions-Koeffizient
FEMALE ROLE	+1,44	DISTRESS	+0.92
SIGN AFFECT	+1,35	UNDERSTATE	+0,80
DEVIATION	+0,80		
COMMUNICATE	+0,75	MOVE	-0,43
SEX THEME	+0,66	THOUGHT FORM	-0,55
		THINK	-0,78
ANGER	-0,16	DEFENSE	-0,96
MESSAGE FORM	-0,94	AFFECTION	-1,70
GUIDE	-1,12		

Abbildung 34

In den sprachlichen Äußerungen des Patienten schlagen sich besonders deutlich die Hinweise auf „weibliche“ Rollen (FEMALE ROLE) und die Hinweise auf affektive Zustände nieder, wie sie in den Wörtern „Sehnsucht“, „Misstrauen“ und „Lob“ zum Ausdruck kommen. Die positive Bedeutung der Kategorie „DEVIATION“, die die sprachlichen Hinweise auf herzbezogene Anfälle, Tachykardien und Zustände der Übelkeit enthält, fügt sich gut dem klinischen Vorverständnis ein. Weiterhin sprachen für „Kastrationsangst“ die Hinweise auf kom-

munikative Beziehung voraussetzende Ausdrücke wie „Antwort“, „Bestätigung“, „Bemerkung“ u.a. mehr. Die Kategorie „SEX SENSE“ ist ebenfalls ein Prädikator für „Kastrationsangst“, was sich auf die Probleme der Selbstbefriedigung und der Sexualängste des Patienten bezieht. Beim Text des Analytikers trägt die Erwähnung des Wortes „Angst“ (DISTRESS) deutlich zur Vorhersage bei; ebenfalls findet sich der deutungstechnisch zu wertende Hinweis auf die Kategorie UNDERSTATE, die praktisch mit dem Wort „Abstrich“ identisch ist. Der hier untersuchte Analytiker verwendet dieses Wort relativ oft als Hinweis auf die übergroße Rolle des Mangels im Erleben des Patienten.

Einen negativen Beitrag zur Regressionsgleichung leisten die Kategorien GUIDE („Hilfe“ und „Schutz“) und formale Ausdrücke für einen Informationstransfer, wie sie in der Kategorie MESSAGE FORM rubriziert sind. Beim Analytiker ist der ausgeprägt hohe Beitrag der Kategorie AFFECTION auffällig. Wenn der Therapeut also über Gefühle der „Liebe“, „Sympathie“ und „Zuneigung“ spricht, dann wurde die Stunde kaum im Sinne der Kastrationsthematik beurteilt.

Diese „Interpretation“ des Ergebnisses kann nur eine vorläufige sein. Sie enthält Hinweise auf sprachliche Zeichen, die nach ihrem statistischen Verhalten besonders eng mit dem klinischen Konzept verbunden sind, bzw. durch ihre Anwesenheit im Verbatim-Protokoll sich in der Wahrnehmung des klinischen Beurteilers im Sinne eines hemmenden Faktors ausgewirkt haben. Die klinische Relevanz dieser Ergebnisse lässt sich in verschiedener Weise prüfen. Zunächst sollte die Betrachtung einiger ausgewählter Stunden ein Verständnis für das Zustandekommen gerade dieser Linearkombination semantischer Merkmale ermöglichen. In einem zweiten Schritt könnte dann die Vorhersagequalität der Linearkombination für neue, bisher klinisch nicht eingestufte Stunden überprüft werden.

4.3.3.2 Vorhersage des klinischen Konzeptes 2 „Trennungsthematik bei libidinöser Übertragung“

Die schrittweise Regressionsanalyse ergab folgende inhaltsanalytische Kategorien von Patient und Analytiker, die ein multiples R von + .94 mit dem klinischen Konzept 2 aufweisen (Abb. 35).

Schritt-Nr.	Variable	Multiples R
1	APPROACH - A	0.46
2	FEMALE ROLE - P	0.54
3	THOUGHT FORM - A	0.62
4	AROUSAL - P	0.66
5	GOOD - P	0.70
6	TOOL - A	0.75
7	ATTACK - A	0.79
8	SIGN STRONG - P	0.81
9	JOB ROLE - A	0.84
10	PEER STATUS - P	0.86
11	ATTACK - P	0.88
12	UNDERSTATE - P	0.89
13	PLEASURE - P	0.90
14	SOCIAL PLACE - A	0.92
15	GUIDE - P	0.93
16	NOT - P	0.94

Abbildung 35

Mit diesen 16 Kategorien von Patient und Analytiker lassen sich 88% der Varianz des klinischen Konzeptes 2 aufklären. Zum Verständnis des Gewichtes der einzelnen Kategorien führen wir im folgenden die jeweiligen Regressionskoeffizienten an: (Abb. 36)

Patientenvariable	Regressions-Koeffizient	Analytikervariable	Regressions-Koeffizient
FEMALE	+ 0.71	JOB ROLE	+ 0.83
AROUSAL	+ 0.62	THOUGHT FORM	+ 0.60
GUIDE	+ 0.46	APPROACH	+ 0.53
NOT	+ 0.26	UNDERSTATE	+ 0.47
		TOOL	+ 0.32
ATTACK	- 0.29		
SIGN STRONG	- 0.57	ATTACK	-0.41
PLEASURE	- 0.66	SOCIAL PLACE	-0.51
PEER STATUS	- 0.67		
GOOD	- 1.31		

Abbildung 36

Bewertung der Ergebnisse:

Beim Patient erscheint wieder als prädiktive Kategorie die FEMALE ROLE - Variable, die auch bereits bei der Kastrationsangst-Analyse auftrat. Wir vermuten aber aufgrund anderer Befunde, dass in den deutlichen „Trennungs“-Stunden die Kategorie „weibliche Rolle“ hauptsächlich von dem Wort „Mutter“ getragen wird. Eine Überprüfung^{*)} der korrelativen Zusammenhänge zwischen Kastrationsangst, prä-ödipaler Anklammerung und den beiden Wörtern „Mädchen“ und „Mutter“ ergaben sich folgende, unsere Hypothese bestätigende Befunde: (Abb. 37)

	Mädchen	Mutter
Kastrationsangst	+ 0.13	- 0.12
prä-ödipale Anklammerung	- 0.04	+ 0.31

Abbildung 37

Positiv schlägt sich weiterhin das Vorhandensein der Kategorie AROUSAL beim Patienten nieder, womit Wörter wie „Aufregung“, „Gefühl“, „Stimmung“ etc. denotiert sind. „Schutz“ und „Hilfe“ (GUIDE) tragen diesmal zum Unterschied positiv bei und unterstützen die inhaltliche Validität dieser Beziehung. Die Kategorie NOT besteht aus den Wörtern „Gegenteil“ und „Mangel“ und fügt sich in diese Deutung gut ein. Der Analytiker trägt durch Hinweise auf den Arzt und die Krankenschwester (auch Schwester) unbewusst doch als Mutterübertra-

^{*)} Die hier verwendeten Häufigkeitsdaten für die beiden Wörter „Mutter“ und „Mädchen“ wurden mit Hilfe eines speziellen Wörterbuches erhoben, indem alle Wörter des Wörterbuches PIAF 4 als eigene Kategorie fungierten. Hierdurch kann der genaue Anteil eines Wortes an einer Kategorie für jede einzelne Stunde bestimmt werden. Diese Kenntnis ist, wie in diesem Fall, für die Interpretation eines Kategorien-Scores von größter Bedeutung.

gungsfigur denkbar, zur fürsorgenden Thematik bei. Die Kategorie THOUGHT FORM (Gedanke, Bedeutung, Idee, Konzept etc.) taucht auch hier wieder als Beitrag des Analytikers auf, allerdings als Verstärkung des Konzeptes. Es wäre von großem Interesse, diesen Unterschied in der Deutungsstrategie noch genauer zu kennzeichnen. Hypothetisch ist denkbar, dass der Analytiker bei der präödpalen Thematik stärker stützend arbeitet und dem Patient mehr Ideen, Gedanken und Phantasien zu vermitteln sucht. Der positive Beitrag der Kategorie APPROACH (Nähe) ist unmittelbar evident. Die Kategorie TOOL enthält alle jene Gegenstände des täglichen Lebens, die den Patient in der Entfaltung seiner Motorik behindern bzw. ihm diesen müssen (z.B. Taxi, Auto, Türe).

Die klinische Beurteilung einer früheren Übertragungseinstellung ist negativ mit dem Auftreten der Kategorie ATTACK sowohl bei Patient wie Analytiker verbunden. Besonders stark wirkt sich die Anwesenheit von Kategorie GOOD (Gutes) auf die Prädiktorfunktion aus.

Hierbei scheint es sich jedoch um einen Artefakt zu handeln. Aus der Korrelationsmatrix geht hervor, dass die Kategorie GOOD mit dem klinischen Konzept 2 nur mit +0.17 korreliert, was bei N = 55 nur auf dem 20%-Niveau rangiert.

Da die Kategorie GOOD auch nur in 13 von 130 Stunden auftritt, basiert diese Korrelation in der Stichprobe von 55 Stunden auf ca. 5 - 6 Stunden.

Der negative Beitrag der Kategorie PLEASURE, mit der die Äußerungen des Patienten über seine Lustlosigkeit erfasst werden, macht deutlich, dass das hier erfasste klinische Konzept die Äußerungen der Lust/Unlustproblematik des Patienten nicht abdeckt. Aus diesem Grunde kann die Kovarianz von Trennungsangstthematik und libidinöser Einstufung, die bei der faktoriellen Untersuchung auf eine gemeinsame Dimension reduziert werden konnte, auch „präödpale Anklammerung“ genannt werden; die lustlosen-aggressiven Aspekte dieser frühen Übertragungseinstellung werden damit nicht erfasst. In diesem Sinne ist die negative Gewichtung der Kategorie PLEASURE eine deutliche Bestätigung der klinischen Überlegungen.

4.3.3.3 Vorhersage des klinischen Konzeptes 3

„Aggressive Übertragung mit diffuser Angst“

Die schrittweise Regressionsanalyse ergab folgende inhaltsanalytische Kategorien von Patient und Analytiker, die zusammen ein multiples R von + 0.77 aufweisen.

Schritt-Nr.	Variable		multiples R
1	ANGER	- P	0.44
2	JOB ROLE	- A	0.55
3	DISTRESS	- P	0.63
4	THINK	- A	0.67
5	MOVE	- A	0.71
6	AROUSAL	- P	0.74
7	DEFENSE	- A	0.77

Abbildung 38

Mit diesen 7 Kategorien von Patient und Analytiker lassen sich 59,3% der Varianz des klinischen Konzeptes erklären. Zum Verständnis der Gewichtung der einzelnen Variablen führen wir im folgenden erneut die Tabelle mit den jeweiligen Regressionskoeffizienten an (Abb. 39).

Patientenvariable Name	Koeffizient	Analytikervariable Name	Koeffizient
DISTRESS	+ 0.71	JOB ROLE	+ 1.08
ANGER	+ 0.57		
		THINK	– 0.55
AROUSAL	– 0.44	DEFENSE	– 0.56
		MOVE	– 0.58

Abbildung 39

Beim Patient erscheint hier die Kombination von Angst (DISTRESS) und Wut (ANGER), wie wir sie bereits in einer früheren Arbeit registrieren konnten. Die Rollen des Arztes und der Krankenschwester (JOB ROLE) werden vom Analytiker offensichtlich in deutlicher Beziehung zur aggressiven Übertragung gesehen.

Hemmend für die Einstufung im Sinne der aggressiven Übertragungskonstellation wirken die Kategorien AROUSAL beim Patienten-Text (so als ob Aufregung, Stimmung, Neugierde zu zahme Ausdrücke für das Erleben des Patienten in diesen Stunden sind); beim Analytiker taucht wieder die deutungstechnische Kategorie „Abwehrmechanismen“ auf (DEFENSE), verbunden mit dem Verzicht auf Hinweise zur Aktivität (MOVE) und Verzicht auf Hinweise zur Erkenntnis, Grund oder Zweifel (THINK).

4.3.4 Zur klinischen Prüfung der Ergebnisse

Die multiple Regressionsanalyse berechnet eine Linearkombination unabhängiger Inhaltskategorien derart, dass die durch die Linearkombination erklärte Varianz der abhängigen klinischen Variablen ein Maximum wird. Nun haben wir bei dieser Anwendung des Regressionsmodelles verschiedene Einschränkungen anzubringen, deren Auswirkung auf die faktische Vorhersage – und Erklärungskraft für die Beurteilung von Behandlungsstunden im Einzelnen zu prüfen ist.

Die hauptsächlichen Einschränkungen, die wir hier aufführen müssen, beziehen sich:

- a) auf die Qualität der klinischen Einstufungen und
- b) auf die Qualität der inhaltsanalytischen Methode.

ad a) Von dem bisher beschriebenen Aufbau der Untersuchung her ist es deutlich, dass die Validität der inhaltsanalytischen Erklärung mit der Gültigkeit der klinischen Einstufungen steht und fällt. Sollten die Kliniker sich bei ihren Beurteilungen auf jeweils wechselnde Sprachinhalte bezogen haben, so müssen sich zwangsläufig Fehlkodierungen mit der inhaltsanalytischen Methode ergeben. Ebenfalls ist die „Nicht-Übereinstimmung“ der Beurteiler, die wir ja nur teilweise durch die Verwendung des faktorenanalytischen Ordnungskonzeptes eliminieren können, eine mögliche Quelle für erhebliche Abweichungen, die im Einzelfall schwer auf den einen oder anderen Beurteiler zurückgeführt werden kann.

Es würde sich anbieten, zu einem späteren Zeitpunkt für jeden Beurteiler getrennte Regressionsanalysen zu berechnen, um diese Fehlerquelle besser abschätzen zu können.

ad b) Die Qualität der inhaltsanalytischen Methode ist insofern zu kritisieren, als die Beschränkung auf Substantive einen erheblichen Anteil möglicher sprachlicher Schlüsselreize, die dem Kliniker zur Verfügung stehen, nicht berücksichtigt werden konnte.

Diese Beschränkung soll in einer weiteren Untersuchung, bei der ein Teil der hier verwendeten Stichprobe, die z.Zt. in vollem Umfang, d.h. den ganzen fließenden Text umfassend, neu erstellt wird, auf ihre Auswirkung hin überprüft werden. Bei der Untersuchung von Angstäußerungen mit der inhaltsanalytischen Methode ergaben sich bereits Hinweise, dass bei der Verwendung aller Textbestandteile auch deutlich andere Inhaltskategorien bei der Operationalisierung der klinischen Konzepte zum Tragen kommen (KÄCHELE, GRÜNZIG und MERGENTHALER, 1976).

Weiterhin muss die grundsätzliche Kritik an der hier verwendeten Form der Inhaltsanalyse erhoben werden, dass es sich um die sogenannte „Kontext-freie-Einwort-Analyse“ (KOFEINA-Methode) handelt, bei der die semantische Information des Kontextes nicht berücksichtigt werden konnte, sowie die syntaktische Information außer acht bleiben musste.

Bei diesen Fehlerquellen handelt es sich um eine, im deutschen Sprachbereich z.Zt. noch nicht eliminierbaren Sachverhalt, da für den deutschen Sprachbereich noch keine weiterführenden Entwicklungen zum maschinellen Sprachverstehen vorliegen.

Zwar liegen im Bereich der linguistischen Datenverarbeitung eine Reihe von umfangreichen Forschungsprogrammen an, bei denen bisher aber die Lösung syntaktischer Dekodierungsstrategien so im Vordergrund stehen, dass für die pragmatischen Zwecke der angewandten Inhaltsanalyse hieraus z.Zt. noch kein Nutzen gezogen werden kann. Nach DIETRICH und KLEIN gibt es z.Zt. Nur ein „einziges existierendes und implementiertes Verfahren, das tatsächlich jeden beliebigen Satz analysiert“, wobei sich die Analyseleistung auf die Erkennung syntaktischer Merkmale beschränkt (1974, S. 92). Hier sind für die Zukunft sicher Entwicklungen zu erwarten, die den derzeitigen Stand der Inhaltsanalyse als ‚primitiv‘ im wahrsten Sinn des Wortes erscheinen lassen.

Diese Einschränkungen machen eine klinische Überprüfung der regressionsanalytisch ermittelten Ergebnisse notwendig. Hierzu betrachten wir im folgenden zunächst einige Stunden im Detail, bei denen sich besonders deutliche Abweichungen zwischen den klinischen und den maschinell ermittelten Werten für die drei Konzepte ergeben. Wir beginnen mit der 4. Stunde, die einen besonders extremen Wert – klinisch wie inhaltsanalytisch – für das Konzept 1 „Kastrationsangst“ erhalten hat. In dieser Stunde findet sich z.B. folgende Passage:

P.: Was mich gerade beschäftigt – ja ich habe vor zwei Jahren den Entschluss gefasst, meine keusche Haltung aufzugeben, und das hat mich einen ziemlichen Kampf gekostet, an dem ich fast zerbrochen bin. Ich hatte damals – das ist möglicherweise eine Folge davon – zum ersten Mal diese heftigen Herzrhythmusstörungen gehabt und solche Extrasyolen, und zwar ziemlich heftige, die ich bis zum Hals herauf gespürt habe oder die sogar mit irgendwelchen Lichterscheinungen in den Augen verbunden waren. Da hab ich natürlich höllische Angst davor.

A.: Ja – .

P.: Bis ich mich da dran gewöhnt hatt. Und diese innere Unruhe und das Aufgewühltsein, das hat sich natürlich noch verstärkt, wie ich das Mädchen, von dem ich Ihnen erzählt habe, kennen gelernt habe, zumal als ich sah, dass ich mit der Annahme, dass sie bereits sexuelle Erfahrungen hatte, recht hatte. Das war für mich damals aus der bisherigen Haltung heraus immer noch ein Schock, obwohl es mir ja eigentlich gelegen kam, sich genau mit meiner Absicht deckte.

A.: Sie meinen, es war trotzdem ein Schock, obwohl das Mädchen Erfahrungen hatte? Oder was meinen Sie?

P.: Ja, nein, ich meine, dass das Mädchen – die Tatsache, dass das Mädchen bereits sexuelle Erfahrungen hatte, war für mich trotzdem ein Schock, obwohl ich dachte, ich hätte mich jetzt zu einer anderen, der gegenteiligen Haltung bekehrt. Aber das ist anscheinend nicht voll und ganz gelungen. Es hat sich ja auch inzwischen rausgestellt, dass es nicht der Fall ist. Der Umbruch war für mich anscheinend doch viel größer, als ich vorher angenommen hatte.

Pause.

Und da hab ich natürlich lange Zeit Angst vor der ersten intimen Begegnung mit Mädchen. Bei dieser dann selbst, hab ich erstaunlicherweise sehr leicht über mich gebracht. Ich hab mich nicht einmal geschämt, mich nackt auszuziehen und das Mädchen nackt zu sehen.

Pause.

Aber in der Folgezeit hatte ich dann auch wieder Angst, einmal natürlich, dass etwas passieren könnte und auf der anderen Seite immer so auf das Herz bezogen, glaub ich wenigstens, dass es irgendwie in der Belastung nicht stand hätte. Irgend so ein Gefühl hatte ich.

Pause.

A.: Der körperlichen Belastung des Verkehrs?

P.: Ja, ja, ja.

A.: Und sie hatten Angst vor der Konzeption.

Diese Passage schildert für jeden Leser unmittelbar verständlich die Ängste des Patienten, die im Zusammenhang mit sexuellen Beziehungen auftreten, wobei der Schwerpunkt der Angst deutlich bei der Angst vor Beschädigung liegt. Dies ist das Thema der ganzen Stunde. Die klinischen Beurteiler haben auch übereinstimmend die Kastrationsthematik in den Mittelpunkt gestellt, wobei sie aber auch deutliche Hinweise auf Schuld- und Beschämungskonflikte im Material der Stunde wahrgenommen haben. Hinweise auf Trennungsthematik fehlen ganz. Kontrovers ist die Beurteilung der aggressiven Übertragung zwischen den Beurteilern (A und C) bzw. B, der die freundliche Übertragung stärker akzentuiert (Abb. 40).

	A	B	C
Lib. Übertragung	6	13	7
Aggress. Übertragung	13	6	14
Trennungsangst	1	1	1
Kastrationsangst	18	19	19
Schuldangst	20	14	23
Beschämungsangst	12	23	13

Abbildung 40 (Bei den hier aufgeführten Werten handelt es sich um Rang-Werte (!) aus den integrierten Rangskalen)

Beim erneuten Lesen des Verbatim-Protokolles findet sich ein Hinweis auf die mögliche „Überschätzung“ der aggressiven Übertragungsthematik durch die Beurteiler A und B. Nachdem der Patient sich sehr ausführlich mit seinen sexuellen Erfahrungen beschäftigt hat, sich über „freche Zärtlichkeiten“ seinerseits geäußert hat, bekommt er plötzlich Angst, der Psychoanalytiker könnte ihm diese verübeln:

P.: Wenn ich mit jemand spreche, achte ich immer genau auf die Nuancen meines Gegenübers. Ich finde, dass Ihre Stimme wieder an Härte gewonnen hat.

Pause.

Das erste Mal, als Sie mir sagten, ich solle mich hinlegen, war sie gegenüber dem Anfang eigentlich etwas streng, etwas brüskierend im Ton, der weich, behutsam - und das ist ein bisschen wieder verschwunden.

Der Analytiker greift diese flüchtige Übertragungsreaktion sofort auf.

A.: Vielleicht liegt es aber auch daran, dass Sie gerade zwei Bemerkungen von mir als störend empfanden, nämlich meine Frage, was Sie mit frechen Zärtlichkeiten meinen..... jedenfalls meine ich, dass es sein könnte, dass Sie meiner Stimme nun mehr Strenge zuschreiben, weil Sie sich mit einem Thema beschäftigen, wo Sie Vorwürfe erwarten.

Diese Passage widerspiegelt eine kurze Verstimmung des Patienten, die vom Psychoanalytiker aber gleich als Externalisierung seines eigenen Gewisses gedeutet wird und damit auch im manifesten Gehalt der Stunde wieder verschwunden ist.

Vergleicht man nun die klinischen und inhaltsanalytischen Werte dieser Stunde, so findet sich folgender Befund (Abb. 41):

		Kliniker	Computer	Residualwert
Klin. Konzept I	Kastrationsangst	+ 3,34	+ 2,60	+0,74
Klin. Konzept II	(Trennungsthematik bei lib. Übertragung)	- 0,57	- 0,88	+ 0,31
Klein. Konzept III	(diffuse Angstthematik bei aggressiver Übertragung)	- 0,06	+ 0,62	- 0,69

Abbildung 41

Die maschinelle Inhaltsanalyse ergibt für das Konzept der Kastrationsangst einen sehr extremen Wert (den zweitgrößten der ganzen 55 Stunden!), doch ist der klinische Einstufungswert noch deutlich extremer. Keine andere Stunde wurde von den Beurteilern so extrem hinsichtlich der Kastrationsangst beurteilt, dass wir nachträglich an der Validität dieser extremen Einstufung zweifeln müssen. Die hieraus resultierende Diskrepanz scheint uns eher artifizieller Natur zu sein.

Beim erneuten Lesen des Behandlungsprotokolls, von dem ich oben einige Auszüge gebracht habe, bestätigt sich die Festlegung der Stunde auf die Kastrationsthematik, aber hinsichtlich des Ausprägungsgrades wäre man sicher zurückhaltender. Hier kann auch ein methodischer Fehler der Einstufungsuntersuchung zum Tragen kommen, da die Stunden nicht in zufälliger Reihenfolge eingestuft wurden. Der Ankerreiz des einzelnen Beurteilers musste sich erst beim Lesen der Protokolle bilden und somit ist mit einer gewissen Modifikation der Beurteilung im Verlauf der Einstufungsprozedur zu rechnen.

Beim klinischen Konzept II findet die maschinelle Einstufung wenig Anhalte für die Trennungsthematik bzw. für die Übertragungsbeziehung. Betrachtet man hierzu die klinischen Einstufungen, so zeigt sich, dass die libidinöse Übertragungskonstellation sprachlich nicht zureichend realisiert ist. Sie wirkt sich auf die maschinelle Einstufung praktisch nicht aus. Es

ist denkbar, dass sich hier die Unterscheidung von Beziehungs- und Inhaltsaspekt (WATZ-LAWICK, 1969) als nützlich erweisen könnte. Bereits bei der Interpretation der inhaltsanalytischen Variablen konnten wir wenig Hinweise auf die Übertragungskonstellation ausfindig machen.

In einer früheren Untersuchung (SCHAUMBURG, KÄCHELE u: THOMÄ, 1973) konnten wir zeigen, dass die Personalpronomina eine signifikante Beziehung mit den Übertragungskonstellationen aufweisen.

Die maschinelle Einstufung des 3. klinischen Konzeptes in dieser Stunde ergibt eine gewisse Hervorhebung der diffusen Angstthematik, die klinisch nicht in gleichem Maße erfasst wurde. Hierbei trägt in dieser Stunde die relativ häufige Erwähnung der körperbezogenen Ängste bei, die vermutlich bei der klinischen Beurteilung der Kastrationsthematik miterfasst wurden.

Zusammenfassend kann zu der Kodierung dieser Stunde aber gesagt werden, dass die maschinelle Erfassung die wesentlichen klinischen Verhältnisse recht zufriedenstellend widerspiegelt.

Ein weiteres Beispiel soll nochmals verdeutlichen, dass die Beziehung zwischen klinischen und inhaltsanalytischen Einstufungen als sich gegenseitig relativierende Größen betrachtet werden können. So weist die Stunde 253 für das klinische Konzept II eine deutliche Diskrepanz von klinischem und inhaltsanalytischem Wert auf:

Stunde 253

Klinischer Wert Y: – 0.83
Berechneter Wert Y: – 1.57
Residualwert R: + 0.74

Der geringe inhaltsanalytische Wert kam dadurch zustande, dass von den prädiktiven Kategorien praktisch keine in dieser Stunde aufgetreten ist. Nun ist es ja denkbar, dass die inhaltsanalytische Messung, die auf der Gesamtheit der klinischen Einstufung von N = 55 beruht, korrekter als die klinische Einstufung dieser einen Stunde ist. Betrachtet man unter diesem Gesichtspunkt die klinische Beurteilung der Stunde durch die drei Rater, so fällt eine deutliche Diskrepanz auf (Abb. 42):

Tabelle:

Klinische Einstufung der Stunde 253

Rater	A	B	C
Libidinöse Übertragung	6	13	1
Trennungsangst	1	1	1

Abbildung 42

Das klinische Konzept wurde ja faktoriell gewonnen und wird durch die Skalen „libidinöse Übertragung“ und „Trennungsangst“ der drei Rater getragen. Es ist deutlich zu sehen, dass die Hervorhebung der libidinösen Übertragung durch Beurteiler B zu einer deutlich höheren klinischen Beurteilung beigetragen haben dürfte, als sie nach Ansicht der beiden anderen Ra-

ter angebracht schien. Wir sind der Meinung, dass – auch nach erneutem Lesen der Stunde – die Inhaltsanalyse hier zu Recht korrigierend wirkte.

In dieser Stunde geht es auf dem Boden einer guten Arbeitsbeziehung um die Besprechung der Beziehung von Angst und Wut, die sich sehr fruchtbar gestaltet. Direkte Zärtlichkeitswünsche von Seiten des Patienten sind aber nicht enthalten, so dass eine Hervorhebung der libidinösen Übertragung auf den Intensitätswert 2 (der in der integrierten Rangskala sich als Zehnerwert niederschlägt) nicht gerechtfertigt erscheint.

Zum Vergleich betrachten wir die Stunde 355 für das klinische Konzept II, die einen besonders negativen Residualwert aufweist:

Stunde 355

Klinischer Wert Y. – 0.40
Berechneter Wert Y: + 0.63
Residualwert R: – 1.03

Die klinische Einstufung der drei Beurteiler weist für die beiden zugrundeliegenden Skalen folgende Werte auf:

Tabelle:

Klinische Einstufung der Stunde 355

Rater	A	B	C
Libidinöse Übertragung	4	8	1
Trennungsangst	3	1	3

Abbildung 43

Die Stunde selbst ist durch lange stützende Deutungen des Analytikers gekennzeichnet, was sich bereits an der Verteilung der Redeaktivität^{*)} von Patient und Analytiker ablesen lässt. 74 Zeilen vom Patient werden in 198 Zeilen des Analytikers eingebettet. Der Analytiker thematisiert das Problem des Patienten, keine Antworten auf seine Fragen zu bekommen, was sich auch bei der Inhaltsanalyse in der häufigen Erwähnung der Kategorie THOUGHT FORM (67%) beim Analytiker-Text widerspiegelt. Beim Patient finden sich die Kategorien AROUSAL und GUIDE: AROUSAL bezieht sich auf die vom Patient geäußerte „miese Stimmung“ und GUIDE bezieht sich auf die Hilfe, die er vom Analytiker erwartet, aber nicht bekommt.

Es ist deutlich, dass die inhaltsanalytische Kodierung jene Elemente entdeckt, die positiv zur stützenden, prä-ödipalen Thematik beitragen. Es fällt allerdings auf, dass nur 6 der insgesamt 16 Prädiktoren im Text angesprochen sind; da aber vorwiegend negative Prädiktoren fehlen, führen die positiven Prädiktoren (AROUSAL, GUIDE) beim Patient und THOUGHT FORM und APPROACH beim Analytiker zu der verhältnismäßig positiven Kodierung. Die klinischen Beurteiler haben offensichtlich die sprachlichen Inhalte hier differenzierter gewürdigt. Das faktische Vorhandensein von Prädiktoren wurde offensichtlich durch andere Aspekte, die der Stunde ihr charakteristisches Flair geben, überlagert. So sieht der Beurteiler A in dieser Stunde hauptsächlich eine diffuse Angst, die er als behandelnder Arzt beim Lesen des Proto-

^{*)} Zur Definition der Redeaktivität und zu ihrer Bestimmung durch die einfache Methode der Zeilenausählung, s. KÄCHELE et al., 1973.

kolles vermutlich rückgeschlossen hat; der Rater C kennzeichnet die Stunde durch eine Beschämungsthematik, der sich der Rater B ein Stückweit anschließt.

Ebenso stellt sich die Frage, ob nicht bei dieser Stunde die extreme Verteilung der Redeaktivität vom klinischen Beurteiler als wichtigerer Prädiktor für das vorherrschende klinische Konzept angesehen wurde. Die Einbeziehung dieser formalen Parameter würde hier u.U. wichtige, bisher vernachlässigte ‚cues‘ aufgreifen.

So fand z.B. SPENCE (1969) bei der Überprüfung eines Konzeptes für die analytische Produktivität des Patienten (SIMON, 1968), dass auch die Länge der Sätze ein wichtiger Prädiktor war. Mit der besonderen Rolle der „minor encoding habits“ für die Bewertung literarischen Materials hat PAISLEY (1964) sich ausführlich beschäftigt. Für den psychotherapeutischen Bereich wurden diese Variablen unseres Wissens bisher nicht berücksichtigt. (Die Verwendung syntaktischer Merkmale findet sich allerdings bei STEINGART und FREEDMAN, 1972.)

4.3.5 Die Prüfung der Generalisierbarkeit der Ergebnisse

Die für die klinischen Konzepte empirisch ermittelten Linearkombinationen der Inhaltsvariablen sind zunächst nur für die Stichprobe gültig, an der die Regressionsanalyse durchgeführt wurde.

„Es versteht sich, dass die so gebildete Vorhersagegleichung kein durch die Daten bestätigtes Modell ist, sondern nur den Rang einer Hypothese hat, die durch die Überprüfung an anderen Datensätzen erst zu bestätigen oder zu falsifizieren ist“ (ALLERBECK, 1972).

Neben der Überprüfung des Modells an neuen Datensätzen gibt es auch die Möglichkeit, zunächst auf statistischem Weg zu prüfen, inwieweit die Stichprobenbefunde auch für die – hypothetische – Gesamtstichprobe gelten, bzw. „inwieweit sie als Schätzungen für Eigenschaften der Variablen in der Population“ Geltung haben. Nach GAENSSLEN und SCHUBÖ (1973) kann dies über ein Prüfverfahren geschehen, bei dem die Anzahl der Personen (hier Stunden der Stichprobe) und die Zahl der ermittelten prädiktiven Variablen mit der Höhe des multiplen Korrelationskoeffizienten in Beziehung gesetzt wird.

Mit Hilfe des Prüfverfahrens „kann festgestellt werden, ob die Wahrscheinlichkeit des in einer Zufallsstichprobe beobachteten Zusammenhanges zwischen der Prädiktorvariablen und den Kriteriumsvariablen.... kleiner oder gleich einem Betrag alpha, dem Signifikanzniveau, ist“ (S. 108).

Signifikanz der multiplen Korrelation

Klinisches Konzept I:	$R = + 0.92,$	$p * 0.01$ bei $R < 0.703$
Klinisches Konzept II:	$R = + 0.94,$	$p * 0.01$ bei $R < 0.703$
Klinisches Konzept III:	$R = + 0.77,$	$p * 0.01$ bei $R < 0.535$

Die statistische Signifikanz der gefundenen Vorhersagegleichungen ist gesichert. Aus verschiedenen Gründen wäre jedoch eine weitergehende klinische Überprüfung der inhaltsanalytischen Voraussagen, an denen wir bereits arbeiten, wünschenswert.

4.3.6 Der Verlauf inhaltsanalytisch ermittelter klinischer Konzepte

Abschließend soll die Vorhersageleistung der inhaltsanalytischen Operationalisierung als beschreibende Kraft für die eingangs genannte Zielvorstellung dargestellt werden: nämlich „Verläufe“ klinischer Konzepte maschinell zu ermitteln und zu beschreiben.

Hierzu wurden die empirisch ermittelten Linearkombinationen auf die gesamte Stichprobe von 130 Stunden angewandt. Mit Hilfe eines speziellen Programmes PREDICT^{*)} wurden die von dem EVA-System für die 130 Stunden ermittelten inhaltsanalytischen Werte in die Regressionsgleichung eingegeben und die Werte für jedes der drei klinischen Konzepte berechnet. Im folgenden geben wir zur Veranschaulichung die graphisch dargestellten „Verläufe“ der drei Konzepte wieder (Abb. 44 a, b, und c).

Verlauf: Klinisches Konzept 1

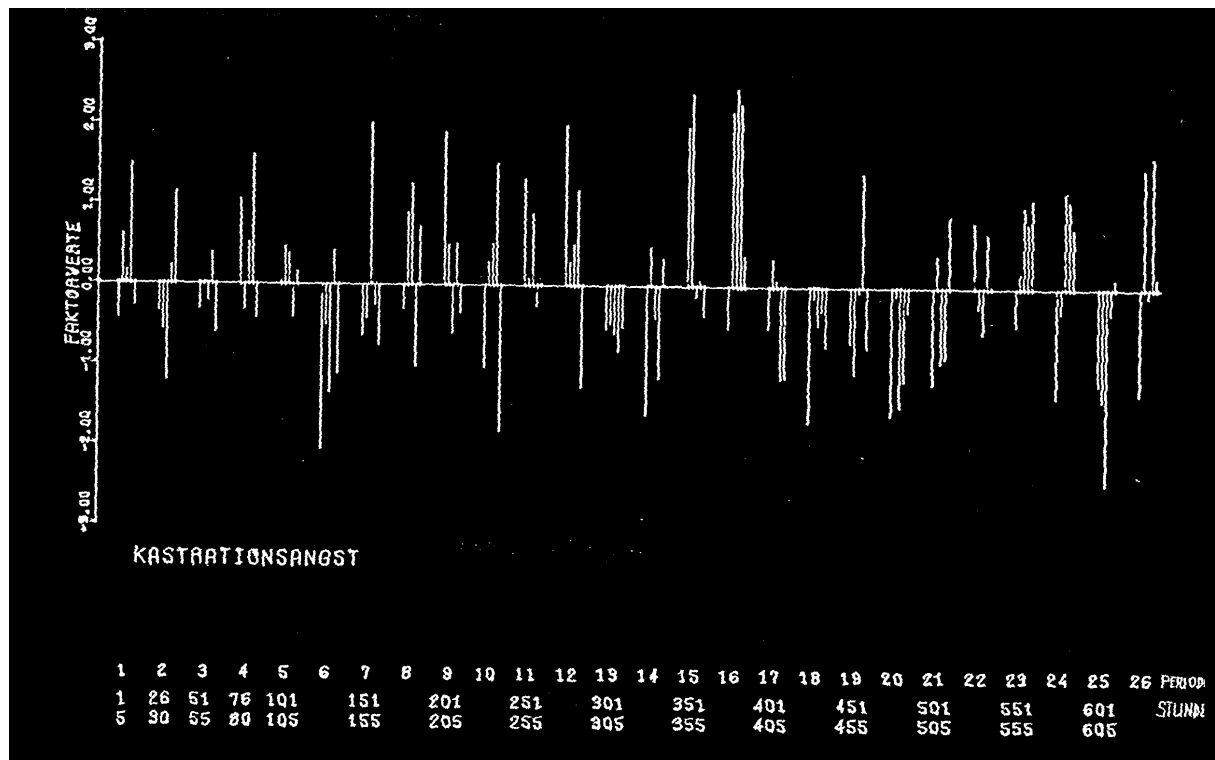


Abbildung 44a

^{*)} Das Programm wurde für diesen Zweck von Herrn Dr. Grünzig geschrieben.

Verlauf: Klinisches Konzept 2

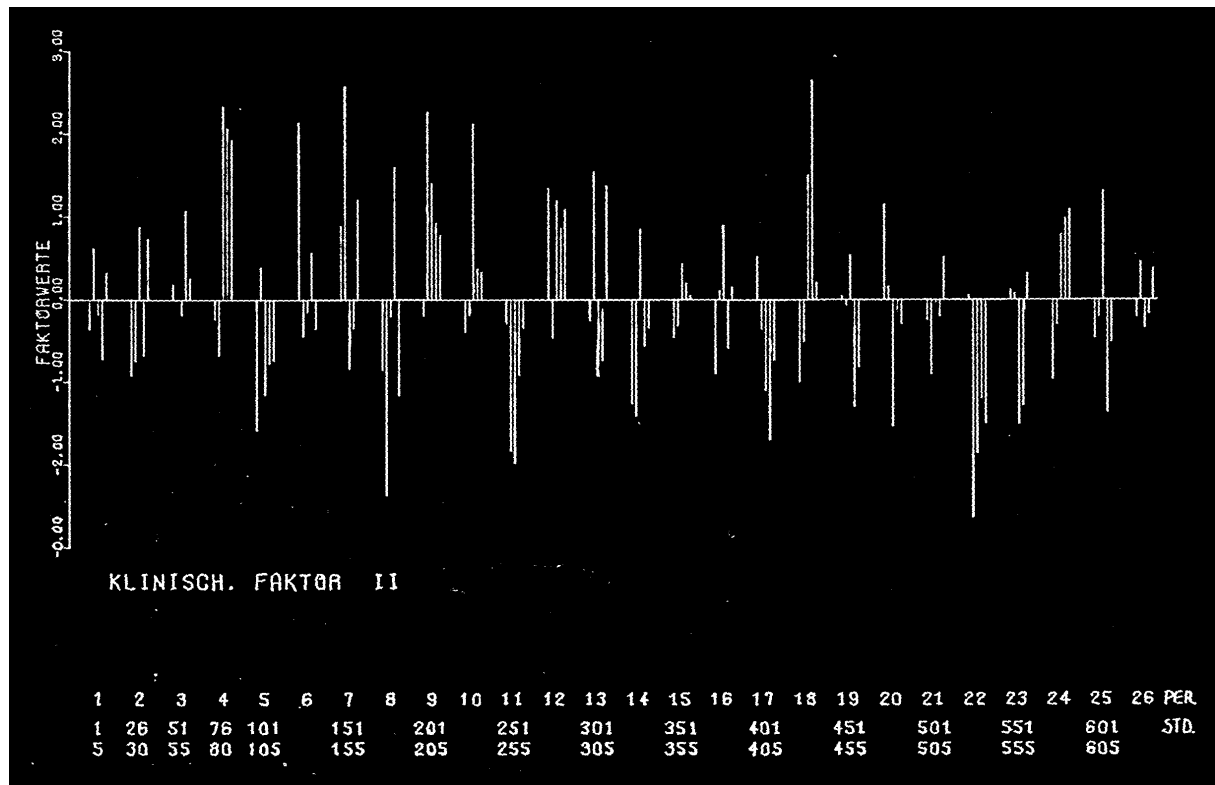


Abbildung 44b

Verlauf: Klinisches Konzept 3

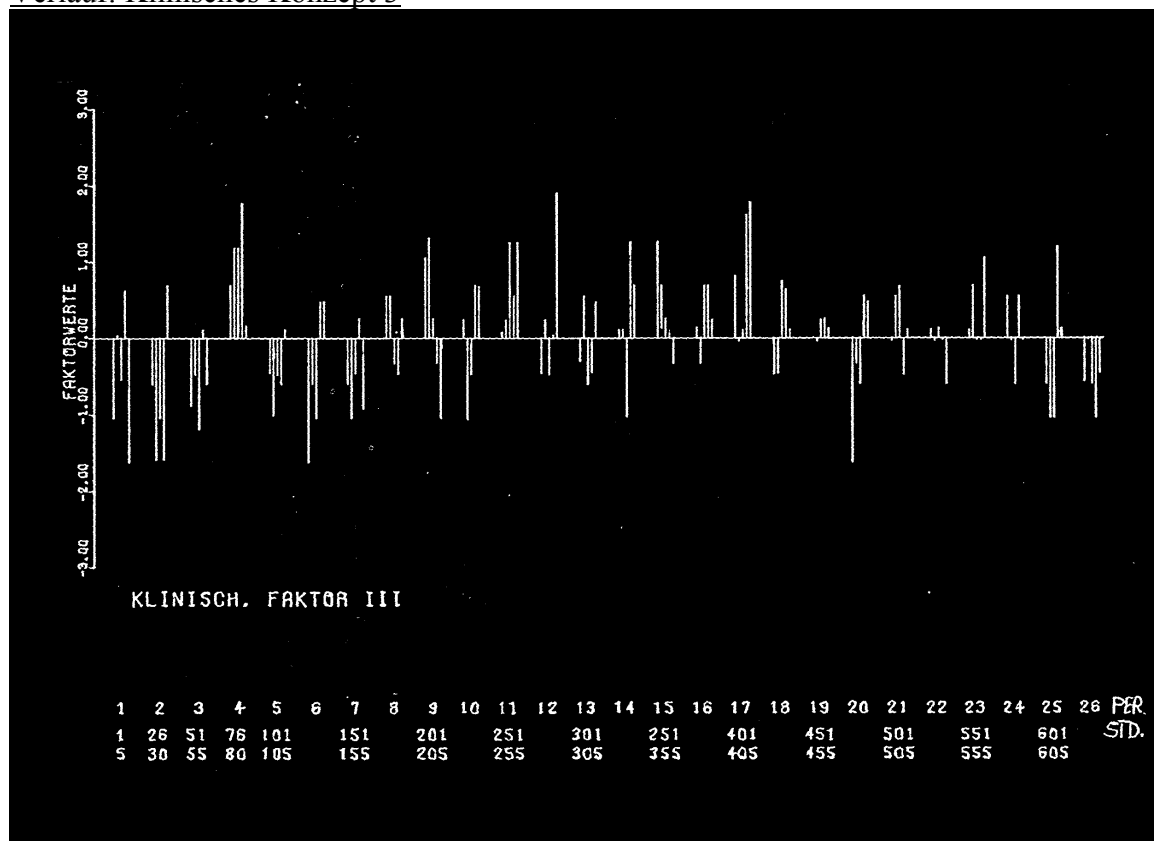


Abbildung 44c

Der Verlauf dieser drei Konzepte, soweit er durch die inhaltsanalytischen Operationalisierungen abgedeckt wird, zeigt keinen offenkundigen Bezug zur Zeit der Behandlung. Ein linearer Zeittrend ließ sich nicht ermitteln. Anders steht es aber mit nicht-linearen Trends, die sich beim Betrachten der Abbildung direkt aufdrängen. Es ist deutlich zu sehen, dass sich Phasen mit eher positiver Ausprägung mit Phasen eher negativer Ausprägung abwechseln. Hier kann nun die weitere Arbeit einsetzen. Die maschinelle Inhaltsanalyse liefert Anhaltspunkte für eine differentielle Typisierung von Behandlungsabschnitten, die im folgenden anhand klinischer Methoden weiter untersucht werden können.

Es würde an dieser Stelle zu weit führen, die verschiedenen Ansätze, an denen wir z. Zt. bereits schon arbeiten, ausführlich darzustellen. Hier ging es darum, zunächst zu zeigen, dass die Mechanisierung der klinischen Einstufungsprozedur durch die Identifizierung des Sprachinhaltes eine Möglichkeit bietet, eine Vielzahl von Behandlungsstunden in den Untersuchungsprozess einzubeziehen, die bei einer rein klinischen Vorgehensweise allein aus Gründen des Arbeitsaufwandes nicht bewältigt werden könnten. Die klinische Einstufung mit ihren Möglichkeiten einer genaueren Differenzierung der Konzepte kann dann wieder ins Spiel kommen, wenn die Inhaltsanalyse eine gewisse Vorklärung eines umfangreichen Materials geleistet hat. So arbeiten wir z. Zt. daran, einen Vergleich der Extremstunden durchzuführen, um auf diese Weise die Operationalisierung der klinischen Konzepte noch weiter zu verfeinern, was dann zu einer erneuten Anwendung führen kann.

4.3.7 Kritische Würdigung der Ergebnisse

Die hier durchgeführte Operationalisierung klinischer Konzepte durch die Methode der multiplen Regressionsanalyse ergab für jedes der drei klinischen Konzepte solche Kombinationen von Inhaltskategorien, die im Ansatz für das Verständnis der Konzepte zutreffend zu sein scheinen. Angesichts des beschränkten Textmaterials – handelt es sich doch nur um Substantive, anstatt vollständiger Sätze – ist jedoch zu erwarten, dass eine Re-Analyse mit vollständigem Textmaterial eine erhebliche Erweiterung und Ausdifferenzierung der Ergebnisse erwarten lässt. Da die Aufbereitung der gleichen Behandlungsstunden mit vollständigem Text z. Zt. Im Gange ist, erschien eine weitere Überprüfung der inhaltlichen Validität der hier erarbeiteten Ergebnisse wenig sinnvoll, da der methodische Aufwand für eine erneute klinische Einstufungsprozedur erheblich ist. Es ging hier zunächst um die Durchführbarkeit solcher Untersuchungen; es ging darum, zu zeigen, dass die Einführung der maschinellen Inhaltsanalyse in die psychoanalytische Prozessforschung offensichtlich keine sinnlosen, uninterpretierbaren Ergebnisse ergibt, sondern im Gegenteil eine überraschende Fülle von Möglichkeiten aufzeigt. Die maschinelle Analyse des Sprachinhaltes von Verbatim-Protokollen stellt nach unseren bisherigen Erfahrungen eine Erweiterung des Arsenal an Forschungstechniken dar, der angesichts des verbalen Charakters der psychoanalytischen Kommunikationssituation eine kaum zu überschätzende Bedeutung zukommen dürfte.