

Grünzig HJ, Kächele H (1978) Zur Differenzierung psychoanalytischer Angstkonzepte. Ein empirischer Beitrag zur automatischen Klassifikation klinischen Materials. *Z Klin Psychol* 7: 1-17

Grünzig HJ, Kächele H

Zur Differenzierung psychoanalytischer Angstkonzepte. Ein empirischer Beitrag zur automatischen Klassifikation klinischen Materials.

Einleitung

Umfassende Prozeßbeschreibungen längerer psychoanalytischer Behandlungen, die in verbatim-transkribierter Form als Datengrundlage vorliegen, scheitern in der Regel an der immensen Fülle der darin enthaltenen Informationen. Es erweist sich stets als äußerst zeitaufwendig, eine hinreichend umfangreiche Stichprobe an Behandlungsmaterial, etwa ganze Behandlungsstunden, Fünf-Minuten-Segmente etc. mit Hilfe von inhaltsanalytisch geprägten Beurteilungsmethoden (z.B. nach Gottschalk und Gleser, 1969; Schöfer, 1977) zu kodieren; neben einem solchen oft nicht tragbaren Zeitaufwand schlagen methodische Vorbehalte, wie unbefriedigende inter- und intrasubjektive Vergleichbarkeit, mögliche Änderungen des Referenzsystems der Beurteilungen im Laufe längerer Beurteilungssequenzen, Ermüdungs- und Monotonierungserscheinungen etc. häufig negativ zu Buche. Durch eine zufriedenstellende Approximierung solcher skalierten Beurteilungen durch geeignete maschinelle, computerunterstützte Methoden könnten diese Nachteile möglicherweise aufgehoben und der personelle Arbeitsaufwand auf ein Minimum reduziert werden.

Durch die in den letzten Jahren beschleunigte Entwicklung von schnellen, in vielfältiger und komfortabler Weise programmierbaren Großrechenanlagen ist es möglich geworden, auch größere Mengen von Verbatimprotokollen mit vertretbarem Aufwand maschinell zu analysieren. Zahlreiche Programmsysteme sind zu diesem Zweck entwickelt worden (vgl. Stone et al., 1966; Gerbner et al., 1969; Deichsel und Holzcheck, 1976). Darüber hinaus gilt es, maschinell analysierbare Parameter zu erarbeiten, die *prozeßsensitiv* und

zugleich *klinisch relevant* sind. Es ist hiermit das recht komplexe Problem der inhaltlichen Validierung solcher maschineller "Beurteilungsergebnisse" angesprochen. Entspricht das Resultat, das sich aus einem formalisierten und programmierten Beurteilungsalgorithmus ergibt, der zur Beurteilung anstehenden klinischen Variablen? Würde ein in Skalierungsaufgaben geschulter, erfahrener Kliniker zu denselben Beurteilungsaufgaben gelangen? Zur Beantwortung dieser und ähnlicher Fragen wollen wir mit den vorliegenden Untersuchungen einen empirischen Beitrag leisten.

Zunächst wollen wir das Problem etwas näher beleuchten, welche Textmerkmale es denn sind, die den klinischen Beurteiler zu seiner Einstufung verleiten. Es sei einmal angesetzt, daß der zu beurteilende Textauschnitt eine eindeutige Beurteilung zuläßt, und daß die Einschätzung "richtig" ist; zu heuristischen Zwecken sei eine solche Formulierung gestattet, auch wenn sie natürlich erkenntnistheoretische Validitätsprobleme in sich birgt. geht man relativ unvoreingenommen, d.h. ohne Berücksichtigung einer expliziten Beurteilungs- und Deutungstheorie an dieses Problem heran, so drängt sich die Überlegung auf, daß die Grundlage für die vorgenommene Beurteilung des Textes eben dieser Text in seinem Sosein ist. Offensichtlich sind es zunächst nur die manifesten Bestandteile des Textes - einzelne Wörter, Ausdrücke, Idiome etc. - welche zu der gegebenen Beurteilung führen. Diese Behauptung wird keineswegs widerlegt, wenn der klinische Beurteiler sein systematisches Theoriewissen einbringt, mit dessen Hilfe von dem primär vorhandenen Textmaterial auf nicht direkt beobachtbare Phänomene und Bedeutungen geschlossen wird; hiermit ist das "Lesen zwischen den Zeilen" mit dem "dritten Ohr" (Th. Reik) angesprochen. Denn auch daraus resultierende mehr oder minder treffende kombinatorisch-interpretative Leistungen können sich ja letztlich auch nur auf das vorgelegte Textmaterial in der gegebenen Formulierung stützen. Auf jeden Fall muß auch die gewagteste und beobachtungsfernste Deutung ihr "fundamentum in re", ihre Verankerung (Cohen, 1969) im gegebenen Oberflächentext aufweisen.

Fragestellung und Untersuchungsansatz

Für eine empirische Untersuchung (Thomä et al., 1976), die sich ausführlich mit der Konsensus-Problematik in der Psychoanalyse beschäftigt, wurden aus der psychoanalytischen Angsttheorie in Anlehnung an Gottschalk und Gleser (1969) vier Angstthematiken abgeleitet und in detaillierten operationalen Definitionen für skalierte Einstufungsuntersuchungen zugänglich gemacht. Es handelt sich um Trennungsangst, Schuldangst, Kastrationsangst und Beschämungsangst. Die operationalen Definitionen sind in der Tabelle 4 (S.) aufgeführt. Aus den Verbatimprotokollen der 500 Sitzungen einer psychoanalytischen Behandlung wurden zunächst 200 Äußerungen des Patienten ausgewählt, welche thematisch diese Angstformen repräsentieren; über Schätzskalen wurden dann für jede der vier Angstformen die 10 prägnantesten ermittelt und für die Untersuchung verwendet (vgl. Thomä et al., 1976, S. 1013 f). Diese 4 x 10 spezifischen Angstäußerungen wurden in zufälliger Aufeinanderfolge einer Reihe Psychoanalytiker aus dem gesamten Bundesgebiet wie auch psychoanalytischen Laien zur Beurteilung vorgelegt, welche der vier Angstformen in der jeweiligen Äußerungen erkennbar sei. Bei dieser Zuordnungsaufgabe wurden insgesamt durchschnittlich 26 von den 40 Äußerungen den jeweils richtigen Angstthematiken zugewiesen; in der Zuordnungsrichtigkeit zeigten die psychoanalytischen Experten und Laien keinen bedeutsamen Unterschied.

Die einleitend aufgeführten Überlegungen sollen nun auf das soeben vorgestellte Material angewandt werden, was im einzelnen zu folgenden Fragestellungen führt:

1. Lassen sich unter Zugrundelegung einer geeigneten Modellvorstellung die den vier Angstthematiken jeweils zugehörigen Äußerungsgruppen allein aus der Kenntnis von Textmerkmalen voneinander trennen? Mit anderen Worten: Ist es möglich, die einzelnen Äußerungen unter ausschließlicher Berücksichtigung von textimmanenten Merkmalen den richtigen Angstthematiken zuzuordnen?

2. Wenn sich solche textuellen Merkmalsdifferenzen als relevant erweisen sollten, wäre zu fragen, welche Textmerkmale es jeweils sind, die den zu gegebenen Differenzierungen führen. Lassen sich inhaltliche Beziehungen dieser Merkmale zu der psychoanalytischen Angsttheorie erkennen, aus der die operationalen Angstdefinitionen abgeleitet und dementsprechend die Auswahl der Äußerungen getroffen worden waren?

3. Lassen sich die aus den bisherigen Überlegungen zu erhaltenden Resultate ebenso erfolgreich auf andere, weniger prägnante Äußerungen des Patienten wie auch des Psychoanalytikers aus derselben psychoanalytischen Behandlung anwenden?

In der Formulierung der Fragestellungen ist bereits der Ansatz einer multivariaten Varianzanalyse (MANOVA) mit einer unabhängigen Variablen, nämlich der Angstthematik in nominaler Skalierung enthalten; das angezeigte Verfahren der Diskriminanzanalyse - als einfache Variante von MANOVA - ermöglicht darüber hinaus eine gewissermaßen "automatische" Zuordnung der einzelnen Äußerungen zu den Angstthematiken aufgrund von Linearkombinationen der abhängigen Variablen. An dieser Stelle wollen wir etwas näher auf die modellhaften Überlegungen eingehen, die diesem statistischen Verfahren zugrunde liegen.

In den genannten Fragestellungen ist implizit das Problem thematisiert, auf welche Art und Weise irgendwelche im Untersuchungsmaterial enthaltenen Zeichen (clinical cues) vom Beurteiler kombinatorisch zu einem Urteil verarbeitet werden. Die logisch vorgeordnete Frage nach der Auswahl und Wahrnehmung solcher Zeichen bleibt hier unberücksichtigt, da in der hier herangezogenen empirischen Untersuchung durch die standardisierte Vorgabe von operationalen Definitionen der Angstkonzepte mit entsprechenden beobachtungsnahen Hinweisen die Möglichkeit der freien Auswahl jeweils für relevant erachteter Zeichen im Textmaterial erheblich eingeschränkt war; somit mußte eine systematische Untersuchung des Prozesses der Wahrnehmung von klinisch bedeutsamen Zeichen entfallen.

Bei der Wahl einer Modellvorstellung für die Kombination von Zeichen entschieden wir uns in Anlehnung an Goldberg (1968) und Cohen (1969) für die einfachste Form, für das linear-additive Modell sich als unzureichend erweisen sollte.¹ Die übliche Formulierung dieses Modells in regressionsanalytischer Notation läßt sich auf unsere Fragestellung nicht anwenden, da hier die abhängige Variable - die Angstform - nominalskaliert ist; in der beschriebenen Untersuchung von Thomä et al. (1976) wurde ja nur die Aufgabe gestellt, eine Äußerung jeweils einer der vier Angstformen zuzuordnen; mit welcher Deutlichkeit etwa die erkannte Angstthematik in der Äußerung enthalten sein mag, wurde in der Untersuchung nicht erfaßt. Die Berücksichtigung der diskriminanzanalytischen Fragestellung erlaubt nun folgende Formalisierung unserer Modellannahme:

$$Y_j = \sum_{i=1}^t c_{ij} X_i + k_j$$

- j Gruppenindex (hier j=1, 2, 3, 4)
t Zahl der Variablen, die berücksichtigt werden
c_{ij} Gewichtungsfaktor der Variablen i in der Gruppe j
X_j beobachteter Wert der Variablen i
k_j Konstante für die Gruppe j.
(vgl. Anderson 1966, S. 172).

Wir setzen also an, daß der Beurteiler die entsprechenden Parameter des Modells impliziert aufgrund theoretischer Vorkenntnisse, klinischer Erfahrung, persönlicher Gleichungen etc. mit zunächst nicht näher bestimmbar aktuellen Werten belegt und im konkreten Beurteilungsprozeß diesem Modell folgt; er ordnet hiernach eine Äußerung derjenigen Angstthematik j zu, für die sich nach dem Modell der höchste **y_j**-Wert ergibt. Die Prüfung der "Richtigkeit"

¹) Eine ausführliche Diskussion des Problems, ob der Kliniker zu seinem Urteil aufgrund linearer oder non-linearer Modellvorstellungen gelangen, hat kürzlich D. Schulz (1975)

dieses Modells erfolgt in erster Linie unter dem Gesichtspunkt, "wie genau es die klinischen Beurteilungsergebnisse vorherzusagen erlaubt" (Goldberg, 1968, S. 485). Ziel der hier vorgelegten Untersuchung ist es also, die Angstthematiken als klinisch beobachtete und quantitativ beurteilte Phänomene inhaltsanalytisch zu beschreiben, nicht als Angstthematik an sich.

Durchführung der Untersuchung

Die quantitative Analyse der Textmerkmale wurde mit maschinell-inhaltsanalytischen Techniken durchgeführt, ähnlich jenen, die im "General Inquirer" (Stone et al., 1966) beschrieben sind. Als inhaltsanalytisches Wörterbuch wurde das "Psychosociological Dictionary Harvard III" in der Ulmer Fassung zugrunde gelegt (s.d. Kächele, 1976); dieses Wörterbuch besteht aus 84 Bedeutungskategorien, die verschiedene Theoriebereiche innerhalb des psychosozialen Forschungsgebietes beinhalten und gewissermaßen das Universum möglicher Textmerkmale konstituieren.²

Entsprechend den Annahmen der quantitativen Inhaltsanalyse (Berelson, 1952; Gerbner et al., 1969; Bessler, 1972; Kächele, 1976) wurde ausgezählt, wie häufig die einzelnen Kategorien, durch entsprechende einzelne Wörter operationalisiert, in jeder der 41 Angstäußerungen aus der Untersuchung von Thomä et al. (1976)³ angesprochen sind. Diese Auszählprozedur wurde mit dem inhaltsanalytischen Programmsystem EVA (Elektronische Verbalanalyse) von K. Holzscheck und E. Mergenthaler (Holzscheck, 1976; Grünzig et al., 1976a) auf der TR 440-Anlage des Rechenzentrums der Universität Ulm vorgenommen.

vorgelegt.

²) Das Wort "Angst", welches nach der Wörterbuchvorschrift in die Inhaltskategorie "DISTRESS" hätte eingeordnet werden müssen, ist als sogenannte Einzelwortkategorie verwendet worden, weil es in dem hier anstehenden Gesamttextmaterial dermaßen häufig aufgetreten ist, daß es die Kategorie "DISTRESS" invalidiert hätte.

³) Den 40 originalen Äußerungen aus der kurz referierten Untersuchung von Thomä et al. (1976) wurde eine weitere Äußerung hinzugefügt, die ebenfalls hoch übereinstimmend der Kastrationsthematik zugeordnet worden war.

Da die Äußerungen aus nur wenigen Zeilen Text bestehen, sind viele der Inhaltskategorien überhaupt nicht oder eher selten angesprochen; wegen der zu geringen Trennungsfähigkeit solcher Kategorien sind diese aus den weiteren quantitativen Untersuchungen eliminiert worden. Eine zunächst durchgeführte univariate Varianzanalyse der verbliebenen 26 Inhaltskategorien ermögliche einen ersten Überblick über die Diskriminanzfähigkeit der einzelnen Textmerkmale. Die anschließend vorgenommene Diskriminanzanalyse gab Auskunft über die Kombination derjenigen Inhaltskategorien, die mit entsprechenden Gewichtungen versehen zu einer maximalen Trennung zwischen den vier Angstgruppen führt. Die daraus abgeleiteten Diskriminanzfunktionen wurden auf eine größere Zahl von Einzeläußerungen angewandt, die ebenfalls nach den vier Angstformen beurteilt worden waren. Die automatische Klassifikation der neuen Äußerungen erfolgte unter Verwendung der Parameter der Diskriminanzfunktionen, die sich aus der Analyse der 41 Äußerungen ergeben hatte.

Auswertung und Ergebnisse

In die weiteren quantitativen Analysen wurden also, wie erwähnt, 26 Inhaltskategorien einbezogen; diese wurden wegen der vorgefundenen Datenstruktur weiterhin danach dichotomisiert, ob die gegebene Kategorie in der gegebenen Äußerung angesprochen wurde oder nicht. Wegen der Kürze der Äußerungstexte wurden nur in wenigen Fällen und auch nur bei wenigen Kategorien Häufigkeiten über 1 erhalten, so daß diese Dichotomisierung keinen wesentlichen Informationsverlust mit sich bringt.

a) Auswertung in univariater Betrachtung (Varianzanalyse)

Entsprechend der Fragestellung 1 wurde für jede der 26 einzelnen substantiellen Inhaltskategorien geprüft, ob zwischen den vier Äußerungsgruppen (zu je 10 Äußerungen; die Gruppe "Kastrationsthematik"

besteht allerdings aus 11 Äußerungen) Unterschiede hinsichtlich der Kategorienmittelwerte bestehen.

Da es sich hier um dichotome Variablen handelt, stimmt diese Art der einfaktoriellen Varianzanalyse recht gut mit der üblicheren Chi-Quadrat-Auswertung überein; das varianzanalytische Verfahren wurde aus Gründen der Modellvorstellung und aus rechen-ökonomischen Gründen gewählt.

In der Tabelle 2 sind die 11 wichtigsten Inhaltskategorien mit ihren Gruppenmittelwerten und den varianzanalytischen F-Werten aufgeführt. Bei der Kennzeichnung des jeweiligen Signifikanzniveaus ist ein möglicher Selektionseffekt noch nicht berücksichtigt; dieser beinhaltet die Überlegung, daß z.B. bei 20 unabhängig vorgenommenen Prüfungen bereits ein auf dem 5 %-Niveau gesichertes Ergebnis erwartet wird.

An dieser Stelle sei bereits festgehalten, daß die Wörterbuch-Kategorien URGE (Sehnsucht, Wunsch, wollen) überwiegend bei Trennungsangst-Äußerungen, IF (oder, vielleicht, wenn) vorwiegend bei Kastrationsangst- und Schuldangstäußerungen, UNDERSTATE (aber) vorwiegend bei Kastrationsäußerungen, SIGN ACCEPT (ja) vorwiegend bei Kastrations- wie auch bei Trennungsangstäußerungen auftreten.

b) Auswertung in multivariater Betrachtung (Diskriminanzanalyse)

In diesem univariaten Untersuchungsansatz wird stets jeweils nur eine einzige Inhaltskategorie als abhängige Variable berücksichtigt. Es stellt sich nun die Frage nach einer multivariaten Analyse, in der für die Diskriminierung zwischen den Angstgruppen mehrere Kategorien zugleich herangezogen werden. Dieser multivariate Ansatz entspricht zweifellos dem klinischen Beurteilungsprozeß wesentlich besser, denn auch der Beurteiler wird sicher nicht nur ein semantisches Merkmal für die Gruppenzuweisung berücksichtigen, sondern mehrere für ihn bedeutungsvolle Kategorien betrachten und sie womöglich noch mit unterschiedlichem Gewicht für seine

Zuordnung heranziehen. Die Formalisierung einer solchen Modellvorstellung haben wir bereits durchgeführt.

Für die im folgenden mit dem Programm BMD07M (Dixon, 1973) durchgeführte schrittweise Diskriminanzanalyse wurde eine Vorauswahl der Variablen vorgenommen; in die Diskriminanzanalyse gingen nur diejenigen 11 Inhaltskategorien ein, deren univariate F-Werte größer als 1 ist (vgl. Tabelle 2). Insgesamt werden die vier Äußerungsgruppen durch diese 11 Inhaltskategorien deutlich voneinander getrennt; Wilk's Lambda beträgt 0.13, was bei $df = 33,37$ auf dem 1 %-Niveau statistisch gesichert ist ($F=2.39$ bei $df = 33,80$); die nur vier Gruppen unterscheiden sich also insgesamt hinsichtlich der Mittelwerte aller 11 Variablen. Allerdings gelingt die paarweise Trennung zwischen den Schuld- und Beschämungsäußerungen, den Schuld- und Kastrationsäußerungen und den Trennungs- und Kastrationsäußerungen nicht in befriedigendem Maße, wie die folgende F-Werte-Tabelle 1 zeigt.

T a b e l l e 1

Mittelwertsunterschiede zwischen je zwei Gruppen bei gleichzeitiger Berücksichtigung aller 11 Variablen

	Beschämungsangst	Kastrationsangst	Schuldangst
Kastrationsangst	3.06		
Schuldangst	1.46	2.01	
Trennungsangst	4.33	1.79	3.06

(Bei $df = 11.27$ ist ein F-Wert ab 2.99 auf dem 1 %-Niveau gesichert)

Ein wesentlicher Grund für die unbefriedigende Trennung der Kastrationsangstäußerungen von den anderen liegt darin, daß zwei Inhaltskategorien, welche aus statistischen Gründen nicht in die Diskriminanzanalyse aufgenommen werden konnten, besonders hohe Mittelwerte in den Kastrationsangstäußerungen aufweisen; offenbar sind es gerade diese Kategorien, die besonders deutlich die Kastrationsthematik von den anderen unterscheiden. Die in dieser Diskriminanzanalyse unbefriedigende Absonderung der Kastrationsangstäußerungen ist folglich als methodisches Artefakt zu werten. Die Diskriminanzfunktionen sind in der folgenden Tabelle 2 aufgeführt.

T a b e l l e 2

Diskriminanzfunktionen.

Die Werte in den Klammern geben die Häufigkeiten an, mit denen die Kategorien in den Äußerungsgruppen auftreten

	Beschämungs- angst	Kastrations- angst	Schuld- angst	Trennungs- angst	F(df)
					=3.37

Variable

TIME REFERENCE

SPACE REFERENCE

DEVIATION

URGE

IF

NOT

WORK

MOVE

UNDERSTATE

SIGN ACCEPT

ANGST

Konstante

* $p < .10$

** $p < .01$

Die aufgrund dieser Diskriminanzfunktionen vorgenommenen automatischen Klassifikationen der Äußerungen führen im einzelnen zu folgenden Zuordnungen:

T a b e l l e 3

Resultate der Klassifikationen aufgrund der Diskriminanzfunktionen

Zahl der Äußerungen, die diesen Gruppen
automatisch zugeordnet worden sind

	<u>BA</u>	<u>KA</u>	<u>SA</u>	<u>TA</u>	<u>Gesamt</u>
Beschämungsangst (BA)	9	0	1	0	10
Kastrationsangst (KA)	0	9	2	0	11
Schuldangst (SA)	1	1	8	0	10
Trennungsangst (TA)	0	1	0	9	10

Die eingangs gestellte Frage (vgl. S. 3), ob ein inhaltsanalytisch geprägtes Modell mit automatischen Zuordnungsmethoden geeignet ist, einen eher klinischen Beurteilungsprozeß im Resultat abzubilden und zu replizieren, kann prinzipiell positiv beantwortet werden: rein formal, d.h. unter Berücksichtigung der Wahrscheinlichkeiten, daß eine bestimmte Äußerung der "richtigen" Angstform zugewiesen wird, gelangt das hier vorgestellte automatische Klassifikationsverfahren zu 35 richtigen Zuordnungen, was einer Trefferquote von 85 Prozent entspricht. Dieses Resultat übertrifft noch dasjenige der psychoanalytischen Experten aus der Untersuchung von Thomä et al. (1976), die 71 Prozent der Äußerungen "richtig" zugeordnet haben.

In der Tabelle 4 sind diese 11 Variablen nach den Angstthematiken geordnet, und es ist angegeben, in welcher Richtung sie jeweils diskriminieren. Der Tabelle 2 sind hierfür diejenigen Kategorien entnommen worden, die aufgrund der Häufigkeit die jeweilige Angstthematik vorherrschend bestimmen und zugleich von den anderen Thematiken abgrenzen, sei es in positiver oder in negativer Richtung. Die in Klammern aufgeführten einzelnen Wörter sind diejenigen, die die jeweilige Inhaltskategorie bei dem hier verwandten Textmaterial im wesentlichen ausmachen.

T a b e l l e 4

Die diskriminierenden Inhaltskategorien, nach Angstthematiken mit ihren operationalen Definitionen geordnet

Beschämungsangst:

Hinweise auf das Gefühl
der Lächerlichkeit,
der Kränkung,
des Ungenügens,
der Scham und
Erniedrigung

WORK (machen) +
ANGST (als Einzelwortkategorie) +
TIME REFERENCE (Anfang, Jahr,
anfangen, dann immer, jedesmal) -
SPACE REFERENCE (ab, auf, da,
darüber, hier über, unter) -
URGE (Sehnsucht, Wunsch, wollen) -
IF (oder, vielleicht, wenn) -
NOT (kein, nicht, nichts, nie) -
MOVE (gehen, werden) -
UNDERSTATE (aber, bißchen) -

Kastrationsangst:

Hinweise auf Verletzung,
psychische Beschädigung
oder Angst vor
Verletzung, weiterhin
Hinweise, die eine
Bedrohung der Potenz
und des Könnens
überhaupt, repräsentieren

TIME REFERENCE (Anfang, Jahr,
anfangen, dann, immer, jedesmal) -
SPACE REFERENCE (ab, auf, da,
darüber, hier, über, unter) +
IF (oder, vielleicht, wenn) +
UNDERSTATE (aber, bißchen) +
DEVIATION (Atemnot, Extrasystole,
Tachykardie, Übelkeit) -
WORK (machen) -

Schuldangst:

Hinweis auf Kritik,
moralische Verurteilung
Verdammung.
Verurteilung, auf Schuld
oder Bedrohung
durch Schuld

DEVIATION (Atemnot, Extrasystole
Tachykardie, Übelkeit) +
IF (oder, vielleicht, wenn) +
NOT (kein, nicht, nichts, nie) -
URGE (Sehnsucht, Wunsch, wollen) -
SIGN ACCEPT (Ja) -

Trennungsangst:

SPACE REFERENCE (ab, auf, da,

Hinweise auf Verlassen	darüber, hier, über, unter) +
Verlassenheit	URGE (Sehnsucht, Wunsch, wollen) +
Einsamkeit, Verstoßen,	MOVE (gehen, werden) +
Verlust von Unterstützung,	SIGN ACCEPT (ja) +
von Liebe oder	ANGST (als Einzelwortkategorie) -
Liebesobjekt oder eine	
damit verbundene	
Bedrohung	

Im Rahmen der vorliegenden Untersuchung, die ja eher als Methodenstudie angelegt ist, soll auf eine Abklärung der Relevanz dieser inhaltsanalytischen Befunde für die psychoanalytische Angsttheorie sowie für das Verhältnis der speziellen Pathologie dieses Patienten verzichtet werden. Hierfür erscheint uns die zugrunde gelegte Datenbasis als zu schmal und weitergehende interpretative Ansätze als zu spekulativ. Wir wollen statt dessen einen speziellen Aspekt der Verwendbarkeit von maschinell-inhaltsanalytischen Methoden für Operationalisierung klinischer Variablen kurz kritisch beleuchten, den wir an anderer Stelle ausführlicher diskutiert haben (Grünzig et al., 1976 b). Am Beispiel der Inhaltskategorie WORK wird deutlich, daß aus der breiten Palette der Einzelwörter, die in dem von uns verwendeten Wörterbuch diese Kategorie ⁴) ausmachen, nur ein einziges Wort, nämlich das Verb "machen", in den hier verwendeten 41 Textausschnitten auftritt. Hieran wird deutlich, daß für eine Interpretation dieser Kategorie WORK die von Stone et al. (1966) angeführte operationale Definition nur in einem sehr eingeschränkten Sinne herangezogen werden darf. Solche Beobachtungen, die wir auch mit anderen Inhaltskategorien bei den von uns verwendeten Verbatimprotokollen von psychoanalytisch-psychotherapeutischen Behandlungen immer wieder machen, lassen es als unumgänglich erscheinen, die Homogenität und damit auch die Validität solcher Inhaltskategorien, wie sie für die maschinelle Inhaltsanalyse benutzt werden, für jeden weiteren Textkorpus stets von neuem zu überprüfen.

Kreuzvalidierung des Beurteilungsmodells

Die weiteren validierenden Untersuchungen wurden an folgendem Textmaterial vorgenommen. In einer Untersuchung von K ä c h e l e et al. (1978) wurden 8 Stunden aus

⁴) Die Inhaltskategorie WORK besteht beispielsweise aus den Einzelwörtern Arbeit, Kauf, Studium, kleben, machen.

einer psychoanalytischen Behandlung - derselben, auf die sich auch die vorliegende Arbeit stützt - derart ausgewählt, daß nach klinisch-skalierten Beurteilungen jeweils zwei Behandlungsstunden in je einer der vier Angstthematiken sehr deutlich und in den anderen jeweils sehr niedrig ausgeprägt sind. Diese 8 Stunden wurden sodann in einzelne Äußerungen zerlegt und von zwei klinisch wie auch einstufigsmäßig erfahrenen Beurteilern danach eingeschätzt, welche der vier Angstformen die jeweils vorherrschende sei. Von den insgesamt 345 Äußerungen, von denen 194 vom Patienten und 151 vom Analytiker stammen, wurden 55 Äußerungen von beiden Beurteilern übereinstimmend derselben Angstthematik zugeordnet.⁵⁾ Nur diese 55 Äußerungen, die sich auf Analytiker und Patient und auf die Angstformen entsprechend der Tabelle 5 verteilen, wurden für die weiteren Untersuchungen verwendet, denn als Außenkriterium für die automatische Klassifikation soll ja das übereinstimmende Urteil erfahrener Beurteiler verwandt werden.

Tabelle 5

Verteilung der Äußerungen der Validierungsstichprobe auf Patient und Analytiker und auf die 4 Angstthematiken

	Beschämungs- angst	Kastrations- angst	Schuld- angst	Trennungs- angst	Gesamt
Zahl der Äußerungen von					
Patient	7	8	6	9	30
Analytiker	1	5	6	13	25
Gesamt	8	13	12	22	55

Die Randsummenverteilungen wie auch die Besetzungshäufigkeiten der Felder sind in keiner Weise statistisch auffällig. Bemerkenswert ist jedoch, daß die Anteile der Patienten-

⁵⁾ Tatsächlich wurden 58 Äußerungen übereinstimmend zugeordnet; drei Äußerungen waren jedoch auch in unserer ersten Äußerungsstichprobe enthalten, so daß diese für die Kreuzvalidierungsuntersuchung eliminiert wurden.

und Analytikeräußerungen das Verhältnis in der Gesamtstichprobe von 345 Äußerungen widerspiegeln; Patient und Analytiker tragen also etwa zu den gleichen Teilen zu den 55 gesamten Äußerungen bei.

Diese 55 Äußerungen wurden ebenfalls einer maschinellen Inhaltsanalyse mit demselben Wörterbuch unterzogen; die daraus resultierenden Kategorienhäufigkeiten wurden wiederum danach dichotomisiert, ob die gegebene Kategorie in der jeweiligen Äußerung aufgetreten ist oder nicht. Auf diesen Datenkörper wurden sodann die Diskriminanzfunktionen (vgl. Tabelle 2) angewandt und den Resultaten entsprechend jeder der 55 Äußerungen einer der vier Angstthematiken nach der linear-additiven Modellvorstellung automatisch zugeordnet. Die folgende Tabelle 6 enthält das Resultat dieser automatischen Klassifikationen im Vergleich mit den Zuordnungen der klinischen Beurteiler; die Werte in den Klammern sind die Erwartungswerte für die Diagonalfelder.

Tabelle 6

Klassifikation der Äußerungen an der Validierungsstichprobe

		Zuordnung durch Beurteiler				
		Beschämungs- angst	Kastrations- angst	Schuld- angst	Trennungs- angst	Gesamt
automatische	BA 6(3.1)	8	2	5	21	
Zuordnungen	KA 1	2(3.1)	4	6	13	
nach den Diskri-	SA 0	2	5(2.0)	2	9	
minanzfunktionen	TA 1	1	1	9(4.8)	12	
Gesamt	8	13	12	22	55	

Wie der Tabelle 6 zu entnehmen ist, hat die automatische Zuordnung zu insgesamt 22 Übereinstimmungen (Summe der Diagonalwerte) mit den klinischen Beurteilern geführt, das sind etwa 40 Prozent richtige Zuordnungen. Die Differenz der Diagonalwerte von ihren Erwartungswerten (Summe = 12.9) ist nach dem eindimensionalen Chi-Quadrat = 11.58; $df = 4$), nach diesem statistischen Prüfverfahren sind die beobachteten Übereinstimmungswerte in der Diagonalen bedeutsam höher als die aus den Randsummenverteilungen errechneten Erwartungswerte für die Diagonalfelder. Hiernach gelangt die automatische Zuordnung der 55 Äußerungen zu den vier Angstformen zu einer Trefferquote, die mit weniger als einprozentiger Irrtumswahrscheinlichkeit von einer zufälligen Trefferleistung abweicht.

Hiermit kann die Validität der diskriminierenden Variablen für Äußerungen dieses einen Patienten wie auch für die seines behandelnden Psychoanalytikers als bestätigt gelten. Allerdings ist kritisch festzuhalten, daß nur 40 % der 55 Äußerungen durch das Klassifikationsverfahren in Übereinstimmung mit den klinischen Beurteilern zugeordnet wurden; eine Trefferquote, die trotz der statistischen Signifikanz weitere Untersuchungen dieser Art an anderen Äußerungsstichproben erforderlich macht.

Diskussion

Ausgangspunkt für die vorliegende Untersuchung war die Überlegung, daß der Kliniker bei der kategorialen Beurteilung von klinischem Textmaterial, z.B. Verbatimprotokollen von Behandlungsstunden etc., primär auf Merkmale zurückgreift, die im Material selbst enthalten sind. Für die in dieser Untersuchung verwandten Äußerungen eines Patienten konnten 11 textimmanente Merkmale herausgearbeitet werden, die eine hinreichend saubere Unterscheidung zwischen den vier untersuchten psychoanalytischen Angstformen ermöglichen.

Ob das jeweils verwendete Modell der Urteilsbildung - in dieser Untersuchung wurde ein linear-additives Urteilsmodell zugrunde gelegt - zutreffend ist, das heißt, ob es den Prozeß der Urteilsbildung, wie er im Kliniker tatsächlich abläuft, rekonstruiert, muß vorläufig eine offene Frage bleiben. Wir sind keineswegs der Ansicht - im Gegensatz beispielsweise zu Spence und Lugo (1972), die im wesentlichen die gleiche Methodik wie wir, wenn auch auf andere Fragestellungen, verwandten -, daß die Zugrundelegung einer solchen Modellvorstellung notwendigerweise impliziert, daß der Kliniker bei der Urteilsbildung tatsächlich so vorgegangen ist; wir neigen vielmehr zu der Auffassung, daß der Kliniker seiner Modellvorstellung entsprechend vorgegangen sein könnte. Den Test des Modells sehen wir also in Übereinstimmung mit Goldberg (1968) darin, wie genau es das Resultat des Beurteilungsvorganges zu replizieren gestattet. "Die Prüfung des Modells besteht darin, wie gut es sich als Abbild des Zustandes des betreffenden Universums eignet, (...) sondern darin, wie gut es die Resultate klinischen Urteilens vorhersagt." (S. 485).

Darüber hinaus muß bei der Formulierung solcher Modellvorstellungen angesetzt werden, daß ihnen allgemeine Gültigkeit zukomme; man muß davon ausgehen, daß alle Beurteiler dasselbe Urteilsmodell zugrunde gelegt haben könnten. In der Formulierung solcher Modelle ist gegenwärtig kein Raum für individuelle oder gar intraindividuelle Schwankungen und Abweichungen vom Modell. Auf unsere Untersuchung bezogen bedeutet dies, daß wir annehmen müssen, daß die zwei klinischen Beurteiler, die sich an unserer Untersuchung beteiligt haben, keine wesentlichen Unterschiede in ihrem Beurteilungsverhalten, d.h. in der Art und Weise, wie sie zu ihrem jeweiligen Urteil gelangen, aufweisen, und daß weiterhin jeder der Beurteiler in seinem Beurteilungsverfahren zumindest für den Zeitraum unserer Untersuchung konstant geblieben ist. Das sind Annahmen, die zumindest im Rahmen der vorliegenden Untersuchung nicht weiter verfolgt werden konnten, die jedoch die Validität der empirischen Befunde beeinträchtigen können.

Mit den vorgegebenen Untersuchungen wollten wir einen empirischen Beitrag zur maschinellen, computergestützten Kodierung von klinisch relevanten inhaltsanalytischen Skalen leisten. Hierbei sind wir von der Überlegung ausgegangen, das Urteil des Klinikern empirisch mit Hilfe entsprechender statistischer Techniken möglichst genau zu beschreiben. Wir bedienen uns einer Variante der multiplen Varianzanalyse, der Diskriminanzanalyse; an anderer Stelle (Kächele, 1976) verwandten wir multiple Regressionsanalysen. Mit einer induktiven Methodik spüren wir diejenigen Textmerkmale auf, die zu einer maximalen Reproduktion klinisch-empirischer Beurteilungsergebnisse führen (vgl. hierzu auch Spence und Lugo, 1972). Ein anderer methodischer Ansatz schlägt sich z.B. in dem Vorgehen von Gottschalk et al. (1975) nieder, den wir als Realisierung einer deduktiven Methodik bezeichnen wollen. Hier werden deduktiv diejenigen Textmerkmale festgelegt, die unter Berücksichtigung klinisch-theoretischer Sachverhalte zu einer Übereinstimmung mit klinischen Beurteilungen führen sollten. Die empirische Überprüfung solcher Merkmalskombinationen erfolgt erst in einem zweiten Schritt anhand konkreten klinischen Textmaterials.

Eine vergleichende Bewertung dieser zwei methodischen Vorgehensweisen zur maschinell-inhaltsanalytischen Operationalisierung klinisch relevanter Variablen ist zur Zeit nicht möglich, da noch zu wenige Erfahrungen aus diesem Forschungsbereich vorliegen. Vielleicht ist es für den gegenwärtigen Kenntnisstand eine zuträgliche Einstellung, die eine Methode zu verwenden und die andere dabei nicht aus den Augen zu verlieren.

Zusammenfassung

In Anbetracht der Probleme, welche die Anwendung von Methoden der skalierten Beurteilung klinischen Materials für eine umfassende Verlaufsforschung mit sich bringt, ziehen wir computergestützte inhaltsanalytische Methoden in die psychotherapeutische Prozeßforschung ein. Bei der skalierten Einschätzung klinischen Textmaterials greift der Beurteiler, so lautet unsere

grundsätzliche Annahme, neben mehr oder weniger dezidierten Beurteilungs- oder Deutungstheorien letztlich auch auf textimmanente Merkmale zurück. Lassen sich solche Textmerkmale erkennen, wenn die Beurteilungsaufgabe entsteht, welcher von vier operational definierten psychoanalytischen Angstthematiken - Beschämungsangst, Kastrationsangst, Schuldangst und Trennungsangst - einzelne ausgewählte Äußerungen eines psychoanalytisch behandelten Patienten zuzuordnen sind? Mit einem maschinellen Kodierungsverfahren (EVA) wird festgestellt, welche Textmerkmale in den einzelnen 41 Äußerungen auftreten. Eine univariate und eine multivariate Varianzanalyse (Diskriminanzanalyse) dieser Merkmale führen zu einer charakteristischen Kombination von 11-Textmerkmalen, die eine statistisch gesicherte ($P < 0.1$) Trennung zwischen den vier Angstthematiken ermöglicht. Eine enge inhaltliche Beziehung dieser Textmerkmale zur psychoanalytischen Angsttheorie läßt sich nicht klar erkennen. Eine Kreuzvalidierung dieser Merkmale führt zwar zu statistisch gesicherten ($P < 0.1$) Klassifikationen; die geringe Zuordnungsgenauigkeit von 40 % jedoch läßt es als unumgänglich erscheinen, weitere empirische Untersuchungen dieser Art durchzuführen, um zu einer präziseren inhaltsanalytischen Differenzierung und Beschreibung von klinisch-psychoanalytischen Angstkonzepten zu gelangen.

Summary

Introducing computer-aided content-analytic methods into psychotherapeutic research, we are trying to overcome some problems of clinical rating methods in the realm of process research. Our principal hypothesis is that in the rating procedure - besides more or less precise theories of judging or interpretation - the judging clinician refers to intrinsic text characteristics. Can those intrinsic characteristics be recognized when individual utterances of a psychoanalytically treated patient are to be judged along four operationally

defined psychoanalytic concepts of anxiety; those of shame, castration, guilt, and separation? Which of the text characteristics occur in the 41 utterances, is found out by an automatic coding procedure (EVA). An univariate and a multivariate analysis of variance leads to 11 text characteristics which - in a specifically weighed combination - established a significant ($P < .01$) discrimination between the four anxiety concepts. The interpretation of these text characteristics according to psychoanalytic anxiety theory is not clear. A cross-validation of these intrinsic characteristics leads to significant ($P < .01$) classification of another sample of utterances. Although statistically significant, the correctness of classification is not higher than 40 per cent, which makes further empirical investigation indispensable in order to arrive at a more precise content analytic discrimination and description of clinical anxiety variables.