

Zur Konstruktvalidität der Holtzman Inkblot Technik (HIT) bei depressiven neurotischen Störungen

Wolfgang Hiller, Erna Duhm

Max-Planck-Institut für Psychiatrie, München, und
Institut für Psychologie, Universität Göttingen

Zusammenfassung: Die Differenzierungskraft der Holtzman Inkblot Technik (HIT) wurde bei einem Vergleich depressiv-neurotischer Patienten und gesunder Kontrollpersonen untersucht. Es fanden sich bei Depressiven weder Beschränkungen im formal-qualitativen Umgang mit dem Reizmaterial (Erfassungsmodus) noch Hinweise auf inhaltlich auffällige Denkprozesse. Farb- und Schattierungsattribute (Erlebismodus) wurden von gesunden Personen zwar sensibler zur Ausschmückung von Deutungen, aber auch vermehrt in impulsiver Form als primäre Determinanten genutzt. Daher ist ein klarer diagnostischer Rückschluß auf Funktionen der Affektkontrolle nicht möglich. Auch für Bewegungsantworten, die selbstsicheres Handeln indizieren sollen, konnten bei Patienten keine Nachteile gesichert werden. Für Konstruktvalidität der HIT in einem Netz anamnestischer, persönlichkeitspsychologischer und sozialer Daten liegt kein ausreichender Anhalt vor. Insgesamt wiesen gestörte Personen nur vereinzelt ungünstigere Testwerte auf. Die Gültigkeit zentraler HIT-Variablen wird insgesamt pessimistisch beurteilt.

Abstract: The power of discrimination between samples of depressed patients and non-depressed controls was examined for the Holtzman Inkblot Technique (HIT). No marked impairments were found in depressives for variables describing the formal handling of the stimulus-material and pathological thought-processes. Aspects of color and shading were used by controls in a more sensitive way for elaborating concepts, but also in an impulsive manner as primary determinants. Therefore no diagnostic conclusions could be drawn for the function of affect-control. For movement-answers, assumed to indicate the ability of mastery-competence, no disadvantages could be demonstrated for patients. There was no clear evidence for construct-validity of the HIT in relation to a broad body of data concerning anamnestic, social and personality aspects. However, subjects with depressive disorders only sporadically produced unfavorable test-scores, leading to a pessimistic general judgement about the validity of major HIT-variables.

1. Einleitung

In der Persönlichkeitsdiagnostik haben Versuche zur Weiterentwicklung projektiver Verfahren ein zwiespältiges Echo gefunden. Zum einen wurde mit konstruktionstechnisch-metrischen Verbesserungen eine objektivierte und hinreichend zuverlässige Anwendbarkeit der Instrumente erreicht, zum anderen ist aber kein markanter «Durchbruch» auf dem Gebiet der Testvalidierung gelungen (etwa Hörmann, 1964, 1982; Spitznagel, 1982). Für die inhaltliche Interpretation der meisten projektiven Testwerte fehlt eine theoretisch und empirisch ausreichende Grundlage, daher ist die vielfach kritische Haltung gegenüber dem diagnostischen Wert solcher Verfahren (etwa Zubin, 1954; Jensen, 1964) nicht ausgeräumt worden.

Besonders deutlich wird dies am Beispiel der «Holtzman Inkblot Technik» (HIT; Holtzman, Thorpe, Swartz & Herron, 1961). Als Nachfolgeverfahren der Rorschach-Technik (RT) gilt sie als eine der bedeutendsten Neuentwicklungen zur diagnostischen Analyse von Formdeutungen (Hartmann & v. Rosenstiel, 1977; Peterson, 1978; Anastasi, 1982; Schmeiser, 1985). Es existieren zwei parallele Testformen mit jeweils 45 Items (Kleckse und klecksähnliche Reizvorlagen). Die Auswertung erfolgt nach präzisen Signierungsregeln (Neumann, 1977) für 22 Aspekte des Antwortverhaltens, den Testvariablen.

Bei Untersuchungen zur Objektivität und Reliabilität der HIT sind wiederholt Werte ermittelt worden, die testtheoretischen Standards (etwa Lienert, 1969) genügen. Die meisten HIT-

Variablen erreichen Übereinstimmungskoeffizienten von größer .70, häufig liegen die Werte sogar über .90 (zusammenfassend Holtzman & Swartz, 1983). Offenbar lassen sich also Personenmerkmale – welcher Art auch immer – zuverlässig erfassen.

Dagegen sind Befunde zur Kriterienvalidität enttäuschend (Hörmann, 1964, 1982; Suinn & Oskamp, 1969; Engel, 1980; Spitznagel, 1982, vgl. auch Holtzman et al., 1961). Zwar wurden wiederholt plausibel interpretierbare Zusammenhänge zu externen Merkmalen ermittelt, mit Validitätskoeffizienten von meist unter .40 blieben sie aber unzureichend. Als Alternative ist die Strategie der Konstruktvalidierung (Cronbach & Meehl, 1955) vorgeschlagen worden, bei dem *nicht* vorrangig nach besonders engen Zusammenhängen zu Außenkriterien gesucht wird, sondern Testwerte in einen theoretisch begründeten Rahmen von Annahmen, Regeln und Hypothesen «verankert» werden (vgl. auch Hörmann). Eine systematische Konstruktvalidierung der HIT ist bislang nicht erfolgt. Gründe dafür liegen möglicherweise in der komplexen, aufwendigen Form dieser Strategie sowie im unklaren theoretischen Bezug der HIT.

Im Zusammenhang mit Persönlichkeitsveränderungen bei depressiven Neurosen soll in der vorliegenden Arbeit ein Beitrag zur Konstruktvalidierung der HIT geleistet werden. Untersucht werden einige ausgewählte Testvariablen, die sowohl in der Rorschach-Literatur (etwa Beck, 1952; Klopfer, Ainsworth, Klopfer & Holt, 1954; Bohm, 1957) als auch bei der üblichen HIT-Auswertung eine zentrale Rolle einnehmen. Die Variablen lassen sich in einem umfassenderen theoretischen Rahmen als Indikatoren für *Ich-Funktionen* (Bellak, Hurvich & Gediman, 1973) interpretieren und können in vier Gruppen eingeteilt werden:

- (1) Den formalen Umgang der Testperson mit dem Reizmaterial im Sinne des Rorschach'schen «Erfassungsmodus» kennzeichnen die HIT-Variablen *L = Lokalisation* (kodierte Größe des berücksichtigten Klecksareals), *FA = Formangemessenheit* (Angemessenheit der Deutung gegenüber den Formen des Klecksreizes), *FD = Formbestimmtheit* (Prägnanz des gedeuteten Konzepts) und *I = Integration* (Verbindung von Reizelementen zu einer größeren Einheit). *L* weist diagnostisch sowohl auf strukturierte, ganzheitliche Reizauffassung als auch auf die Fähigkeit zur Detailanalyse hin, *FA* gilt als Indikator für

adäquate Reizwahrnehmung und Konzeptbildung, *FD* reflektiert die Fähigkeit zu präzisiertem, differenzierendem Denken und Erinnern und *I* wird als Zeichen integrativen, analytischen Denkens angesehen. Die vier Variablen erfassen Aspekte des Leistungsverhaltens wie Aufmerksamkeit, Wahrnehmung, Denken, Lernen, Erinnern, sprachliche und motorische Intentionalität und können dem von Bellak et al. (1973) geprägten Begriff *autonomer (intellektueller) Funktionen* zugeordnet werden.

- (2) Eine eher inhaltsbezogene Auseinandersetzung mit den Klecksstimuli wird in der HIT-Variablen *V = Abwegige Verbalisation* erfaßt (kodierte etwa fabulisierte Kombinationen, Kontaminationen, Antworten mit inkohärenter, bizarrer, realitätsferner Ausgestaltung). Derartige *inhaltlich und formal auffällige Denkprozesse* haben Bellak et al. (1973) bei ihrer Charakterisierung von Ich-Funktionen zu einem eigenständigen Bereich zusammengefaßt.
- (3) Reaktionen auf Aspekte des Reizmaterials werden durch die HIT-Variablen *C = Farbe* (Determiniertheit einer Antwort durch Farbattribute) und *Sh = Schattierung* (Determiniertheit durch Schattierung) erfaßt. In Anlehnung an den von Rorschach beschriebenen «Erlebismodus» sollen impulsive Persönlichkeiten durch eine starke Reagibilität auf Farbe und Schattierung bei gleichzeitiger Vernachlässigung von Formaspekten gekennzeichnet sein. Dies kann auch als Fähigkeit zu *Trieb- und Affektkontrolle* (Bellak et al., 1973) im Sinne einer Ich-Funktion beschrieben werden.
- (4) Bewegt gesehene Bildern (HIT-Variable *M = Bewegung*) soll nach breit akzeptierter Auffassung eine Projektion eigener Personenmerkmale zugrunde liegen (etwa Hill, 1972). Differenzierend sollen Menschenbewegungen (Subvariable *HM*) positive, stützende Kräfte der eigenen Person anzeigen, Tierbewegungen (*AM*) eher triebhafte Impulse, Objektbewegungen (*m*) dagegen innere Spannungen und Konflikte. Die Fähigkeit zur Deutung von Menschenbewegungen läßt sich der von Bellak et al. (1973) beschriebenen Ich-Funktion der *Handlungskompetenz* zuordnen, die durch die Fähigkeit zu einer effektiven Interaktion mit der Umwelt (Bewältigung) sowie durch Selbstvertrauen charakterisiert ist.

Bei depressiven neurotischen Störungen liegt eine Dysregulation psychischer Prozesse vor, die mit Funktionsstörungen in den vier genannten Bereichen verbunden sein können. Bei Personen mit neurotischen Störungen haben Bellak et al. (1973) unter Einsatz eines halbstandardisierten klinischen Interviews multiple Ich-Funktionsbeeinträchtigungen ermittelt. Speziell bei der Beschreibung depressiver Symptomatik wird auf Einschränkungen des Leistungsvermögens hingewiesen, etwa auf Konzentrations- und Wahrnehmungsmängel, verlangsamtes Denken, vermindertes Reaktionsvermögen bei Lei-

stungsaufgaben (speziell Miller, 1975) sowie auf herabgesetzte Kreativität (vgl. zusammenfassend etwa Tölle, 1982). Depressive Denkstörungen werden in der psychiatrischen Literatur als wahnähnliche Vorstellungen (Wertlosigkeit, Schuld, Sündhaftigkeit, hypochondrische Ängste) und in der kognitiv orientierten Psychologie als kognitive Verzerrungen charakterisiert (etwa Salzman, Goldstein, Atkins & Babigan, 1966; Beck, 1972; Arieti & Bemporad, 1978; Tölle, 1982). Einschränkungen der Fähigkeit zu sicherem, kompetentem Handeln können bei Depressiven ein vielfältiges Bild annehmen, etwa durch die Neigung, Anforderungen auszuweichen und vor Schwierigkeiten zu resignieren, durch Lähmungen des Antriebs, Kraftlosigkeit, Willens- und Entscheidungsschwäche (Dilling & Rabin, 1967; Tölle, 1982).

2. Untersuchungsmethode und Hypothesen

Mit der HIT (Individualform A) wurden $n = 35$ Patienten der psychiatrischen Klinik des Max-Planck-Instituts für Psychiatrie in München mit einem depressiven Syndrom getestet. Es handelte sich um zehn Männer und 25 Frauen mit einem mittleren Alter von 37,6 Jahren. Die Untersuchung erfolgte innerhalb der ersten vier Tage nach Aufnahme in die Klinik. Alle Patienten erfüllten die Kriterien einer «Major Depression» nach DSM-III (Köhler & Saß, 1984). Die Kriterien wurden im Anschluß an die HIT-Untersuchung mit einer deutschen Version des von Robins, Helzer, Croughan & Ratcliff (1981) entwickelten «Diagnostic Interview Schedule» (DIS; Wittchen & Rupp, 1981) erhoben. Patienten mit bipolarer oder schizoaffektiver Psychose waren ausgeschlossen. Ebenfalls aus $n = 35$ Personen bestand eine in ihrer Struktur (Alter, Geschlecht, soziale Herkunft, Schulausbildung) mit der depressiven Gruppe vergleichbare Kontrollgruppe. Keine Vergleichsperson war zum Zeitpunkt der Untersuchung nach den DSM-III-Kriterien für «Major Depression» akut depressiv.

Für einen Vergleich beider Gruppen wurde erwartet, daß sich bei Depressiven Beeinträchtigungen von Ich-Funktionen in allen genannten HIT-Variablen durch einen erhöhten Anteil ungünstiger Testwerte niederschlagen müßten.

Zum Zwecke einer umfassenden Konstruktvalidierung wurden ferner externe Daten (Verhaltensweisen, Selbsteinschätzungen, biographische und anamnestische Angaben, Symptome) erhoben, die typischerweise mit depressiven Störungen einhergehen oder ihnen folgen. Dies erfolgte über Selbsteinschätzungen mit Hilfe des Gießen-Tests (GT; Beckmann & Richter, 1975) und des Biographischen Inventars zur Diagnose von Verhaltensstörungen (BIV; Jäger, Lischer, Münster & Ritz, 1976), ferner über Fremdeinschätzungen durch das «Diagnostic Interview Schedule» (DIS). Die Fragebögen wurden von allen getesteten Personen im Abschluß an die HIT-Untersuchung ausgefüllt.

Mit Depressionen einhergehende Veränderungen wurden für 31 Items und sechs Skalen des GT sowie für 91 Items und acht Skalen des BIV prognostiziert. Dies ist in Tab. 1 zusammengefaßt.

3. Ergebnisse

Die Werteverteilung der intelligenzbezogenen Variablen L, FA, FD und I ist graphisch mit sog. «boxplots» (Emerson & Strenio, 1983) in Abb. 1 dargestellt. Jeweils für Kontroll- (KG) und Patientengruppe (PG) ist der Streuungsbereich von Rohwerten ersichtlich. Innerhalb der Kästen befinden sich 50% aller Werte, «Ausreißer» sind durch Kreuze gekennzeichnet.

Depressive Patienten gaben durchweg weniger großflächige Deutungen und mehr Detail- und Kleindetailantworten als Kontrollpersonen (höhere L-Scores; Median PG 21; KG 14). In der vorliegenden Untersuchung interessierten aber einseitige, unausgewogene Tendenzen zu groß- oder kleinflächigen Deutungen, die sich in erhöhten oder verminderten L-Werten widerspiegeln. Dies kann über Nominalskalentests geprüft werden, wenn mittlere L-Scores als «günstig», nach oben und unten abweichende Werte dagegen als «ungünstig» eingestuft werden.

Der «cut-off» lag bei Werten, die in der KG in weniger als 20 Prozent oder in mehr als 80 Prozent aller Fälle erreicht wurden. Nach diesem Kriterium wurden L-Werte unterhalb von 9 und oberhalb von 30 als «ungünstig» klassifiziert. Wie aus Zeile a in Tab. 2 hervorgeht, sind Zahl


Tabelle 1: Übersicht über Fragebogenitems und -skalen, für die eine Relevanz bei der Erfassung und Beschreibung depressiver Störungen angenommen wird.

Werteverminderungen werden für folgende GT-Items und -Skalen erwartet: 4 (eher unabhängig), 6 (Ärger eher herauslassend), Items 16 (beliebt), 23 (geachtet), 24 (bequem), 25 (anderen nahe), 27 (an schönem Aussehen interessiert), 31 (eigensinnig), 33 (durchsetzungsfähig), 37 (anziehend), 39 (fähig zum Ausgelassen-sein), Skalen 1 (Soziale Resonanz), 2 (Dominanz), 3 (Kontrolle), 6 (Soziale Potenz).

Werteanstiege werden für folgende GT-Items und -Skalen erwartet: Items 2 (ungesellig), 3 (gern sich unterordnend), 5 (stark zur Selbstreflektion neigend), 7 (kaum konkurrierend), 8 (sehr ängstlich), 9 (in der Arbeit kritisiert), 10 (eher mißtrauisch), 11 (Liebesbedürfnisse zurückhaltend), 14 (häufig bedrückt), 15 (eher wenig preisgebend), 17 (kaum fähig zu Dauerbindung), 19 (verschlossen), 22 (selten in Auseinandersetzungen), 26 (phantasiereich), 29 (sehr selbstkritisch), 30 (wenig hingabefähig), 31 (eigensinnig), 34 (in der Liebe wenig erlebnisfähig), 35 (unbegabt zum Schauspielen), 38 (unstetig), 40 (im heterosexuellen Kontakt befangen), Skalen 4 (Grundstimmung) und 5 (Durchlässigkeit).

Für BIV-Skalen wird im folgenden angegeben, ob bei Depressiven Werteerhöhungen (+) oder -verminderungen (-) erwartet werden, für BIV-Items beziehen sich die Prognosen auf eine Zunahme von a- oder b-Markierungen: Skala Familiäre Situation (FAM, +) mit den Items 8 (ein Elternteil dominierend, b), 13 (Eltern wirklichkeitsfremd, a), 15 (Eltern leicht kränkbar, a), 27 (Außenkontakte nicht gefördert, b), 29 (Eltern klagsam, b), 31 (Eltern eher ernst, b), 35 (Eltern eher pessimistisch, b), 43 (Eltern unbeherrscht, b), 54 (Eltern schwer belastbar, a), 64 (wenig Achtung für Eltern, b), 70 (Eltern ablehnend, a), 77 (Eltern nicht einander zugewandt, b), 78 (Ehe eher krisenhaft, a), 83 (Eltern schauspielerhaft, a), 88 (häufig Streit der Eltern, b). Skala Erziehungsverhalten (ERZIEH, +) mit den Items 18 (wenig Vertrauen zu Mutter, b), 21 (inkonsistente Erziehung, a), 30 (Eltern verständnislos, a), 32 (beeinflussende Eltern, a), 34 (kaum Anerkennung durch Eltern, a), 38 (wenig Beachtung durch Eltern, b), 42 (Eltern mit wenig Geduld, a), 48 (anale Themen für Eltern tabu, a), 52 (überfürsorgliche Eltern, b), 55 (unangemessene Leistungsanforderungen, b), 56 (wenige Erklärungen durch Eltern, b), 59 (strafende Eltern, a), 60 (wenig Förderung von Autonomie, b), 73 (schlagende Eltern, a), 81 (wenig Zeit von Eltern, b), 82 ("Zuckerbrot und Peitsche", a), 84 (wenig Interesse von Vater, b), 86 (wenig Freiheit bei Entscheidungen, b), 87 (Geschwister bevorzugt, a), 93 (bei sexuellen Themen verlegen, b). Skala Soziale Lage (SOZLAG, +) mit den Items 11 (Außenseiter in Beruf, b), 12 (fühlt sich nicht anerkannt, b), 20 (kann schwer zuhören, b), 26 (keine gute sexuelle Partner-

schaft, b), 45 (mit Beruf zufrieden, a), 57 (Übelkeit und Erbrechen, a), 66 (kann sich schwer öffnen, b). Skala Psychophysische Konstitution (PSYKON, +) mit den Items 2 (befürchtet Herzstillstand, a), 6 (erschöpft, b), 17 (Unwohlsein), a), 19 (Nervosität vor Prüfungen, a), 28 (Sensibilitätsstörungen, b), 47 (befürchtet Versagen in Beruf, a), 51 (Schwindelgefühle, b), 62 (Angst um Arbeitsplatz, a), 63 (fürchtet Anforderungen, a), 71 (Probleme bei Arbeitsplatzwechsel, a), 79 (wetterfühlilig, a), 96 (schreckhaft, b). Skala Soziale Aktivitäten (SOZAKT, +) mit den Items 1 (heterosexuell zurückhaltend, a), 7 (redet nicht gerne vor anderen, a), 16 (erlangt schwer Zuneigung, b), 33 (eher verschlossen, b), 36 (kann schwer über Probleme reden, b), 61 (kommt schwer ins Gespräch, a), 67 (in Gesprächen zurückhaltend, b), 92 (ergreift wenig Initiative, a), 94 (zurückhaltend mit eigener Meinung, b), 95 (meidet Meinungsverschiedenheiten, b). Skala Ichstärke (ICHSTK, +) mit den Items 3 (mißerfolgsorientiert: Beruf, a), 14 mißerfolgsorientiert: allgemein, a), 23 (rasch gekränkt, b), 25 (kritikempfindlich, a), 41 (bei Konkurrenz empfindlich, a), 49 (bei Zeitdruck empfindlich, a), 58 (wenig Zuversicht bei Leistung, a), 72 (sexuelle Hemmungen, b), 74 (wenig durchsetzungsfähig, b), 80 (verletzbar, b). Skala Neurotizismus (N, +) mit dem Items 4 (nervös, a), 39 (Zukunftsängste, a), 40 (innere Spannungen, a), 44 (Schuldgefühle, a), 53 (gerät leicht aus der Fassung, b), 75 (Minderwertigkeitsgefühle, a), 76 (grüblerisch, a), 85 (leicht gekränkt, a), 89 (stimmungslabil, a), 91 (Herzklopfen, a). Skala Extraversion (E, -) mit dem Items 37 (für ruhig gehalten, b), 46 (spielt anderen nicht gerne Streiche, b), 65 (geht nicht gerne aus, b), 68 (lebt lieber ohne Aufregung, b), 69 (behutsam und zurückhaltend, b), 97 (bevorzugt Alleinesein, b).

 DEPRESSIVE PATIENTEN (n = 35)

 KONTROLLGRUPPE (n = 35)

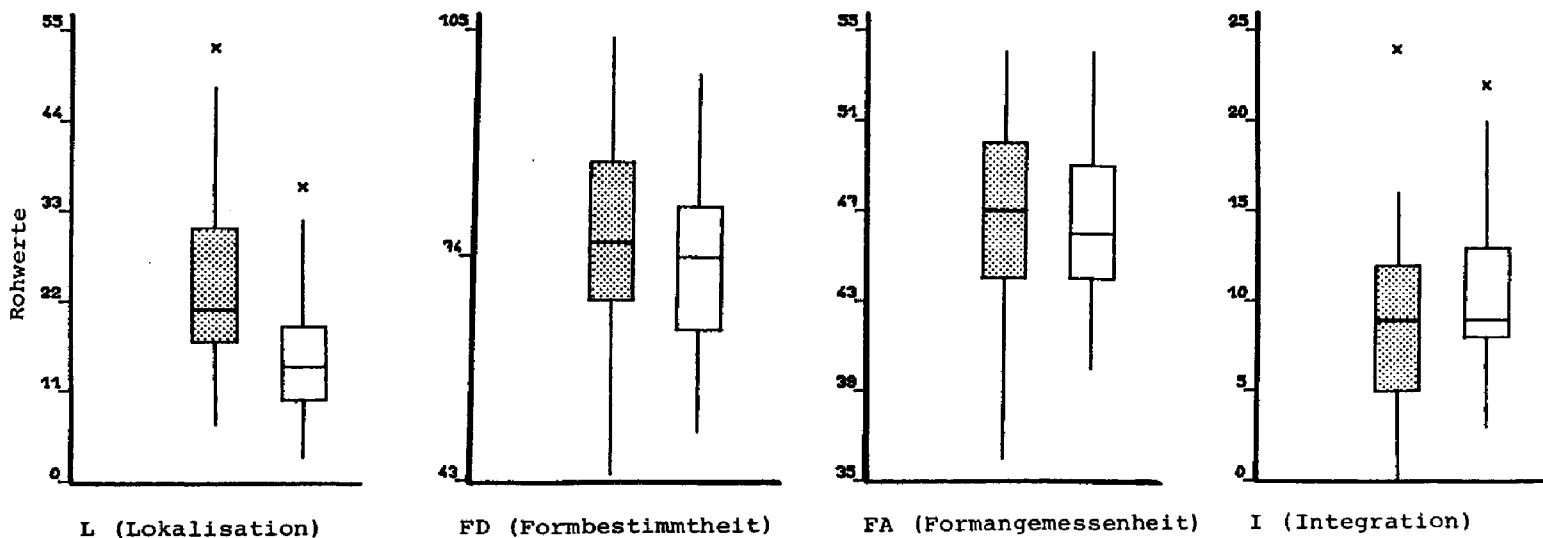


Abbildung 1: Übersicht über die Verteilung von Rohwerten der HIT-Variablen L, FD, FA und I mit «boxplots». Innerhalb der Kästen befinden sich 50% aller Werte, gemeinsam mit den Linien markieren sie den Wertebereich. «Ausreißer»-Werte sind durch Kreuze (x) gekennzeichnet. Die mittlere Linie gibt jeweils den Median an.

Tabelle 2: Verteilung günstiger und ungünstiger Werte in Kontroll- (KG) und Patientengruppe (PG) für intelligenzrelevante HIT-Variablen (Erläuterungen im Text). Angegeben ist für jede der vier Zellen: Zahl der Fälle, prozentuale Häufigkeit innerhalb von PG bzw. KG (in Klammern) sowie empirische Chi-Quadrat-Werte (Chi-Qu).

	KG		PG		Chi-Qu.
	günstig	ungünstig	günstig	ungünstig	
(a)	29 (83%)	6 (17%)	25 (71%)	10 (29%)	0.73
(b)	31 (89%)	4 (11%)	25 (71%)	10 (29%)	2.23
(c)	13 (37%)	22 (63%)	11 (31%)	24 (69%)	0.06
(d)	28 (80%)	7 (20%)	33 (94%)	2 (6%)	--
(e)	27 (77%)	8 (23%)	27 (77%)	8 (23%)	--
(f)	27 (77%)	8 (23%)	20 (57%)	15 (43%)	2.33
(g)	14 (40%)	21 (60%)	12 (34%)	23 (66%)	0.06

und Anteil vorteilhafter L-Scores bei Kontrollpersonen zwar erhöht, der Unterschied zur Patientengruppe (Differenz 12 %) blieb aber statistisch insignifikant. Selbst bei Patienten überwogen unauffällige gegenüber abweichenden Testwerten deutlich (25 zu 10). An diesem Bild änderte sich auch dann nichts, wenn vorwiegend großflächige Antworten (niedrige L-Scores) bei gutem Formniveau (unauffällige oder ausreichende Werte in den Variablen FA, FD, I und V) ebenfalls als «günstig» angesehen wurden (Zeile b in Tab. 2) oder wenn mittlere L-Scores bei Mängeln im Formniveau (abweichende Werte in mindestens einer der Variablen FD, FA, I oder V) ebenfalls das Etikett «ungünstig» erhielten (Zeile c). Bei den Analysen unter b und c wurde also bei der Bewertung von L-Scores *zusätzlich* berücksichtigt, ob Antworten überwiegend präzise und gut strukturiert waren.

Eine Dichotomisierung erfolgte ebenfalls für die anderen intelligenzbezogenen Variablen, wobei Mängel im formalen Umgang mit dem Reizmaterial einseitig durch niedrige Scores angezeigt wurden. Kriterium war jeweils das 20. Perzentil, so daß FD Werte kleiner 62, FA kleiner 44 und I kleiner 8 als «ungünstig» eingestuft wurden. Wie Zeile d in Tab. 2 zeigt, war die Formbestimmtheit (FD) von Klecksdeutungen bei Depressiven besser als bei gesunden Personen. Das geht auch aus der graphisch dargestellten Werteverteilung in Abb. 1 hervor (Median PG 76; KG 75). In der Formangemessenheit (FA) der Antworten unterschieden sich beide Gruppen in den Häufigkeiten dichotomisierter Werte nicht (Zeile e in Tab. 2). Die Mediane beider Gruppen weichen nur geringfügig von

einander ab (PG 47; KG 46), obwohl die Scores bei Patienten im unteren Wertebereich stärker streuen (Abb. 1). Das Integrationsniveau (I) war bei gesunden Personen insignifikant erhöht (Zeile f in Tab. 2), was sich auch in einer gewissen Massierung mittlerer I-Werte (Abb. 1), nicht jedoch in den Medianen zeigte (in beiden Gruppen 9).

Bei einer kombinierten Berücksichtigung der vier intelligenzbezogenen Variablen (Zeile g in Tab. 2) wurden Protokolle nur dann als günstig kategorisiert, wenn durchschnittliche L-Werte sowie keine unterdurchschnittlichen Werte in FD, Fa und I vorliegen (nach den bereits genannten Kriterien). In beiden Gruppen überwiegen in diesem Fall «ungünstige» Protokolle, eine Differenz von 6 Prozent zugunsten der Kontrollpersonen war statistisch insignifikant. Insgesamt ist erkennbar, daß Depressive bei der formalen Handhabung des Testmaterials keine Mängel zeigten, in der Tendenz den Vergleichspersonen teilweise sogar überlegen waren. Es wurden keine intellektuellen Defizite angezeigt.

Für den Bereich der Denkprozesse gaben Kontrollpersonen – entgegen den Erwartungen – deutlich mehr Antworten mit auffälliger Ausgestaltung, erreichten also höhere V-Werte (vgl. Abb. 2). Der Median in der PG betrug 8, in der KG 17. Die Variable V vermag Denkstörungen, wie sie bei Depressiven vermutet werden, offenbar nicht zu erfassen.

Bei den Variablen C und Sh werden Antworten mit der Kodierung «1» als günstig angesehen, da eine kreative Anreicherung der Deutung durch Farb- oder Schattierungsattribute erfolgt, ohne Formaspekte zu vernachlässigen

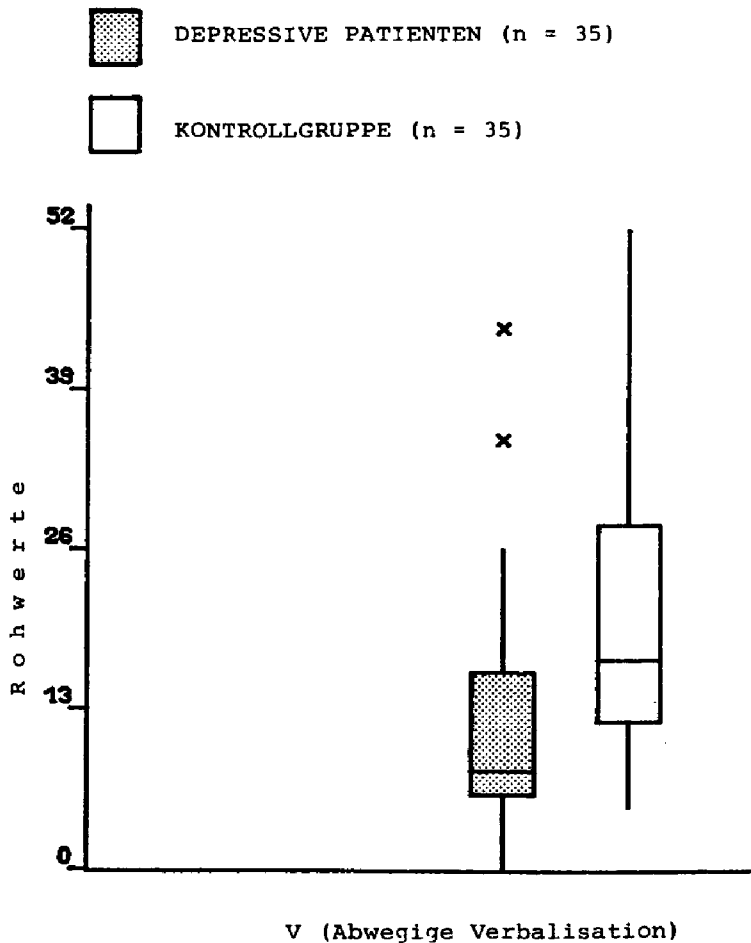


Abbildung 2: Verteilung von Rohwerten der HIT-Variablen V (Abwegige Verbalisation) in den untersuchten Gruppen.

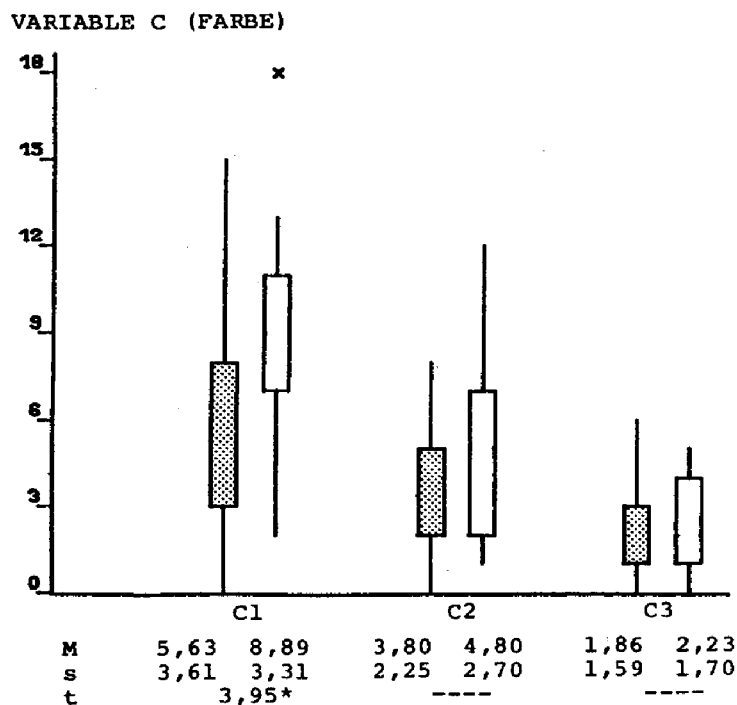
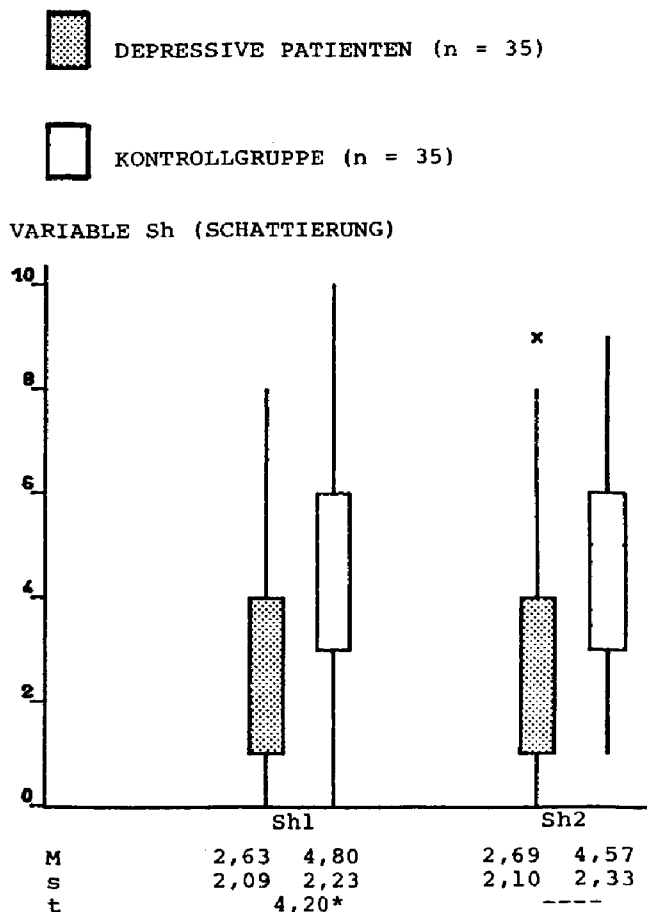


Abbildung 3: Verteilung von Rohwerten für Subkategorien der HIT-Variablen C (Farbe) und Sh (Schattierung). Für Patienten- und Kontrollgruppe sind Mittelwerte (M), Streuungen (s) und bei Abweichungen in der erwarteten Richtung Ergebnisse eines Signifikanztests (t) angegeben.

* = statistisch signifikant bei Alpha = 0.05

(vgl. Neumann, 1977). Auf Impulsivität und Kontrollverlust wird dagegen bei allen *hauptsächlich* farb- oder schattierungs determinierten Antworten («2» und «3» bei C, «2» bei Sh) geschlossen. In Abb. 3 sind Werteverteilungen dargestellt und Mittelwerte für alle Subkategorien von C und Sh angegeben.

Antworten vom Typ C1 und Sh1 traten erwartungsgemäß in der KG signifikant häufiger auf, wie eine Überprüfung der Gruppendifferenzen durch t-Tests ergab. Aber auch Antworten mit höherkodierten Werten, die Mängel in der Kontrolle von Trieben und Affekten anzeigen sollen, wurden in den untersuchten Stichproben von gesunden Personen häufiger gegeben als von Depressiven. Dies widerspricht den Untersuchungshypothesen. C1-Antworten wurden in beiden Gruppen am häufigsten gegeben, C3-Antworten am seltensten. In der Relation zwischen «kontrollierten» und «impulsiven» Ant-



worten (Verteilung von Quotienten in Abb. 4) zeigen sich zwar für die Variable C signifikante Vorteile bei Kontrollpersonen, nicht jedoch für die Variable Sh. Wegen nicht vergleichbarer Intervalle von Quotienten-Werten größer bzw. kleiner als eins erfolgte der Signifikanztest hier mit dem Mann-Whitney-U-Test zur Überprüfung von Rangdifferenzen (etwa Siegel, 1976).

Gleiche Ergebnisse traten ein, wenn die Relation zwischen «kontrollierten» und «impulsiven» C- und Sh-Antworten nicht durch Quotienten, sondern durch Differenzwerte ausgedrückt wurde. Insgesamt ließ sich die Annahme, daß Farb- und Schattierungsantworten affektive Kontrollfunktionen erfassen, nur vereinzelt für die mit «1» kodierten Subkategorien sowie für das Ver-

hältnis günstiger zwischen unterschiedlichen C-Antworten bestätigen.

Handlungskompetenz sollte sich in vermehrten Antworten mit Menschenbewegungen (HM) ausdrücken, ferner in einem günstigen Verhältnis zwischen Menschen- und anderen Bewegungsantworten (AM, m sowie Gesamt-M). Wie aus Abb. 5 hervorgeht, war bei Depressiven die mittlere Anzahl von HM-Antworten zwar erwartungsgemäß vermindert. Entsprechende Nachteile waren aber bei den Relationen nicht mehr erkennbar. In den Quotienten zeigten sich bei depressiven Personen in der Stichprobe sogar günstigere Werte.

HM kovariiert positiv mit der Gesamtvariablen M (Produktmomentkorrelation $r = .74$). Aber ähnlich verhalten sich auch die als eher ungünstig bewerteten Subvariablen AM ($r = .76$) und m ($r = .60$). Ein Werteanstieg für AM und m bei steigendem M-Score ist mit einem entsprechenden HM-Anstieg vergleichbar, erhöhte M-Scores bei gesunden Personen sind also nicht vorwiegend durch einen isolierten oder überproportionalen Zuwachs von Antworten mit menschlicher Bewegung begründet. Die vielfach vorgeschlagene Interpretation von HM-Deutungen als Ausdruck einer stabilen, handlungsfähigen Persönlichkeit ist daher nach dem vorliegenden Befund zweifelhaft.

Zusammenhänge zu den mit den Zusatzinstrumenten GT, BIV und DIS gemessenen Außenkriterien wurden für sieben HIT-Indizes hergestellt: (1) auffällige Neigung zu Ganz- oder Kleindetailantworten (L kleiner 9 oder größer 30), (2) gering formbestimmte Antworten (FD kleiner 62), (3) gering formangemessene Antworten (FA kleiner 44), (4) Antworten mit unterdurchschnittlichem Integrationsniveau (I kleiner 8), (5) auffällige Denkprozesse (V größer 30), (6) ungünstiges Verhältnis zwischen «kontrollierten» und «unkontrollierten» Farbantworten, ausgedrückt im *Quotienten* $Q = C1 / (C2 + C3)$ mit Werten unter dem 20. Perzentil (entspricht Q kleiner 0.833 oder 1:1.2), (7) geringe Anzahl von Menschenbewegungs-Antworten (HM unter 20. Perzentil; entspricht HM kleiner 5).

Wies ein HIT-Protokoll eines der Merkmale auf, so wurde es als ungünstig, andernfalls als günstig eingestuft. Diese Dichotomisierung erfolgte einzeln für alle sieben HIT-Indizes. Für

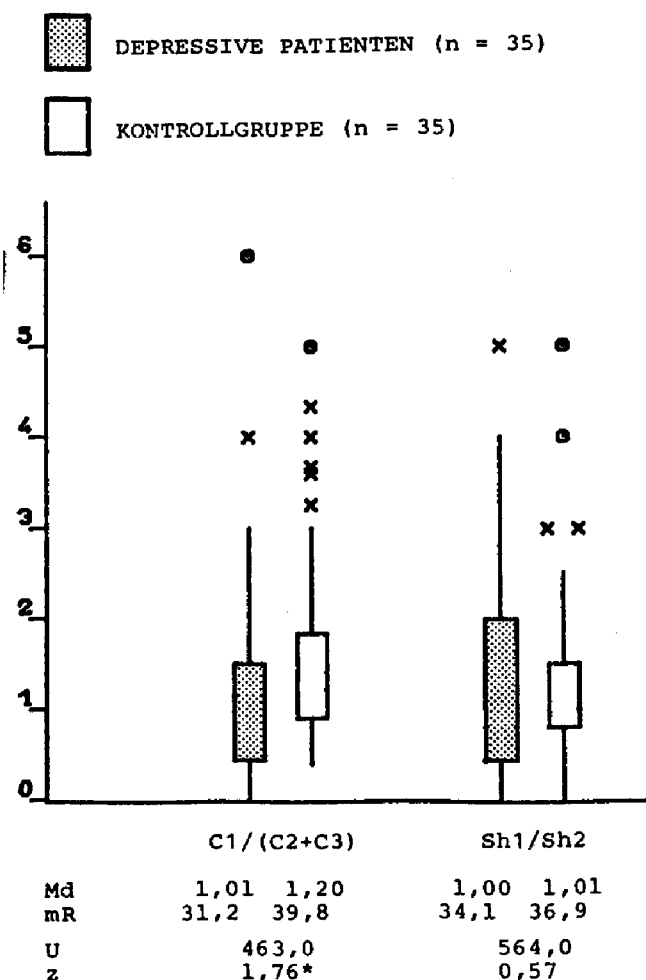


Abbildung 4: Werteverteilung für die Quotienten zwischen «günstigen» und «ungünstigen» Antworten in den HIT-Variablen C (Farbe) und Sh (Schattierung). Gruppenunterschiede wurden mit Hilfe des Mann-Whitney-U-Tests für ordinalskalierte Daten auf Signifikanz getestet. Angegeben sind Median (Md), mittlerer Prozentrangplatz (mR) und z für ermittelte U-Werte.

* = statistisch signifikant bei Alpha = 0.05

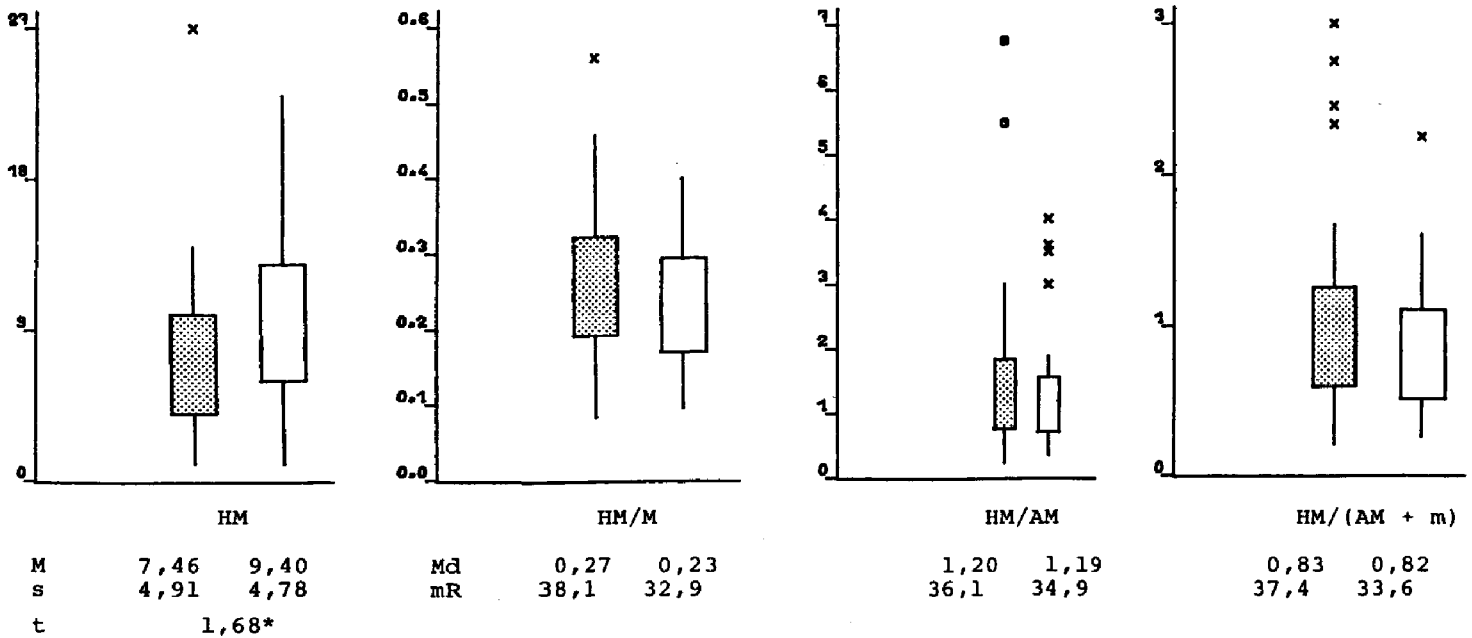
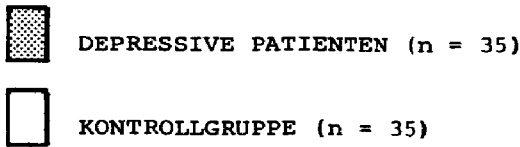


Abbildung 5: Übersicht über die Verteilung von Menschenbewegungsantworten (HM) sowie Werten der Quotienten HM/M, HM/AM und HM/ (AM + m). Nähere Erläuterungen befinden sich im Text. Eine statistische Überprüfung erfolgte für HM mit einem t-Test, für die Quotienten mit dem Mann-Whitney-U-Test. Da für keinen der Quotienten eine Gruppendifferenz in der erwarteten Richtung eintrat, sind U- und z-Werte nicht wiedergegeben.

* = signifikant bei Alpha = 0.05

37 GT-Werte (s. Tab. 1), die in Form abgestufter Selbsteinschätzungen bzw. als Skalenscores vorlagen, wurden mit t-Tests Mittelwertsunterschiede zwischen Gruppen mit günstigen bzw. ungünstigen HIT-Indizes geprüft. In Tab. 3 ist angegeben, bei welchen GT-Items und -Skalen ein statistisch signifikantes Resultat gemäß der

Erwartung eintrat. In Spalte X ist die Anzahl der pro HIT-Index bestätigten Hypothesen angegeben und ihr Prozentanteil an der Zahl der insgesamt durchgeführten Tests (100% = 37 für jeden Index). Unter Y ist aufgeführt, wie häufig ein der Prognose entgegengesetztes Resultat auftrat.

Tabelle 3: Angaben über statistisch signifikante Resultate in erwarteter Form zwischen HIT-Testmerkmalen (HIT) sowie GT-Items und -Skalen. für jedes HIT-Merkmal werden 37 Tests durchgeführt. Erläuterungen im Text.

HIT	Signifikante Resultate:										-Skalen		X	(%)	Y	(%)		
	GT-Items und ...																	
L	22	23	29	33	39							--		5	14	6	16	
FD	-	-										--		0	0	21	57	
FA	35											Sk2		2	5	15	41	
I	3	7	9	22	26	29	31	34	35	37			Sk1	Sk2	12	32	6	16
V	23											--		1	3	28	76	
Q	3	5	6	14	24	26	33	35					Sk2	Sk4	10	27	11	30
HM	2	3	11	24	26	27	30	34	35	37			Sk1	Sk2	12	32	9	24
(a)	3	4	5	6	7	8	9	10	14	23			Sk1	Sk4	23	62	4	11
	24	25	26	29	33	35	37	38	39	40			Sk5					

Bei ungünstig ausgeprägten I-, Q- und HM-Werten weichen Fragebogenscores am häufigsten in der erwarteten Richtung ab, nämlich in knapp einem Drittel der Fälle. Für L werden nur 14% aller Tests signifikant. Keine oder fast keine Erwartung wird bei FD, FA und V bestätigt. Die meisten GT-Items weisen nur bei einem einzigen HIT-Index signifikante Differenzen auf (22, 23, 24, 29, 33, 34, 37). Am häufigsten sind Abweichungen bei den Items 3 und 26 (je dreimal). Das heißt inhaltlich, daß sich Personen mit ungünstigen HIT-Werten vermehrt als von anderen Menschen gelenkt empfinden (3) und weniger Phantasie zu entwickeln glauben (26). Viermal differenziert Item 35 (FA, I, Q, HM), Personen mit ungünstigen HIT-Indizes sprechen sich also häufig «sehr schlechte schauspielerische Fähigkeiten» zu. Bei GT-Skalen treten dreimal Mittelwertsunterschiede für Skala 1 (zunehmend geringere soziale Resonanz bei ungünstigen HIT-Merkmalen) und viermal für Skala 2 auf (vermehrt submissives, nichtdominantes Verhalten).

Wesentlich höher ist die Zahl bestätigter Hypothesen, wenn GT-Scoredifferenzen zwischen den beiden Untersuchungsgruppen (PG vs. KG) getestet werden. Dies ist in Tab. 3 in der Zeile angegeben. 23 Hypothesen sind bestätigt, das sind etwa doppelt so viele wie bei HIT-Indizes mit der stärksten Differenzierungskraft. Der Gießen-Test konnte mit 62 Prozent seiner Items und Skalen zwischen depressiven und gesunden Personen unterscheiden. Der Vergleich macht

deutlich, daß selbst für I, Q und HM eine quantitativ ausreichende Einordnung im Kontext erlebungs- und verhaltensspezifischer Besonderheiten nicht deutlich wird.

Im BIV wurden depressionstypische Veränderungen für 91 Items und 8 Skalen erwartet (s. Tab. 1). BIV-Items sehen die Antwortalternativen a und b vor. Ihre Häufigkeiten wurden für Personen mit günstigen bzw. ungünstigen HIT-Merkmalen ermittelt (erneut einzeln für die sieben bereits genannten Indizes). Bei Abweichungen in der erwarteten Richtung erfolgten Signifikanztests durch bivariate Chi-Quadrat-Verfahren. Für die quantitativ abgestuften BIV-Skalenwerte wurden Gruppenunterschiede mit t-Tests überprüft. Einen Überblick über die Ergebnisse gibt Tab. 4.

Mehr als fünf signifikante Ergebnisse ergaben sich bei keinem HIT-Index. Bei fast 100 Tests und einer Fehlerwahrscheinlichkeit von Alpha = .05 liegt dieser Befund im Zufallsbereich. Bei FD, FA und V bestätigt sich keine einzige Hypothese. Zeile a zeigt, daß der Befund nicht auf mangelnde Differenzierungsfähigkeit des BIV zurückgeführt werden kann, denn bei Trennung der Stichprobe in Depressive und Kontrollpersonen ergeben sich 52 signifikante Resultate. Bei insgesamt 99 Hypothesen ist dies ein erwartungsgemäßer Befund in mehr als jedem zweiten Fall.

Testwerteveränderungen im Widerspruch zu den Prognosen (Spalte Y in Tab. 4) treten bei FD, FA und V in 54 bis 88 Fällen auf. Ungünstig-

Tabelle 4: Angaben über statistisch signifikante Resultate in erwarteter Form zwischen HIT-Testmerkmalen (HIT) sowie BIV-Items und -Skalen. Erläuterungen im Text.

HIT	Signifikante Resultate:										-Skalen			X	Y
	BIV-Items und ...														
L	27	47										FAM SOZLAG PSYKON	5	11	
FD	-	-										--	0	78	
FA	-	-										--	0	54	
I	65	84	94									SOZAKT E	5	37	
V	-	-										--	0	88	
Q	17	61										SOZAKT E	4	42	
HM	61	63										SOZLAG SOZAKT	4	37	
(a)	3	4	6	7	11	12	17	19	20			FAM ERZIEH SOZLAG	52	4	
	28	31	36	39-42	44	47	49					PSYKON SOZAKT			
	51-53	56-59	62	62	65	66						ICHSTK N			
	71	73-80	88	89	91	92	94	96							

ge HIT-Werte sind hier häufiger mit vorteilhaften BIV-Antworten assoziiert als mit störungsrelevanten Aspekten. Inhaltlich ist bei den signifikanten Beziehungen zwischen HIT-Indizes und BIV wenig Übereinstimmung erkennbar. Nur in der Skala «Soziale Aktivitäten» (SOZAKT) treten dreimal bei ungünstigen HIT-Werten erhöhte Mittelwerte auf (I, Q, HM). Einen verminderten Extraversions-Score (E) haben sowohl Personen mit ungünstigen I- als auch mit auffälligen Q-Werten. Soziale Belastungen (Skala SOZLAG) sind vermehrt sowohl mit ungewöhnlichen L-Werten als auch mit einer geringen Zahl von HM-Antworten assoziiert. Auf Itemebene zeigt sich nur bei Item 61 zweimal ein signifikantes Resultat: Sowohl Personen mit ungünstigem Farbantworten-Quotient (Q) als auch Personen mit wenigen HM-Antworten geben – im Vergleich zu Personen mit günstigen Scores in diesen beiden Indizes – häufiger an, mit fremden Menschen nur schwer ins Gespräch zu kommen.

Mit dem DIS-Interview wurde eingeschätzt, ob eine akute depressive, psychiatrisch bedeutsame Symptomatik vorliegt. Berücksichtigt wurden Kriterien der DSM-III-Diagnose «Major Depression» (Köhler & Saß, 1984), nämlich (a) Appetit- und Gewichtsstörungen, (b) Schlafstörungen, (c) psychomotorische Beeinträchtigungen (Verlangsamung oder Agitiertheit), (d) Interesseverlust, (e) Vitalitätsstörungen, (f) eingeschränkte Konzentrations- und Denkfähigkeit, (g) Selbstabwertungen sowie (h) Suizidalität. Es wurde für jedes Symptom ein gehäuftes Auftreten ungünstiger HIT-Indizes erwartet. Ergebnisse einfacher Vierfelder-Tests (Chi-Quadrat) sind in Tab. 5 zusammengefaßt.

HIT	Signifikant: DIS-Symptomgruppen	X		Y	
		(%)	(%)	(%)	(%)
L	--	0	0	1	13
FD	--	0	0	7	88
FA	--	0	0	6	75
I	--	0	0	0	0
V	--	0	0	8	100
Q	M, V, S	3	38	0	0
HM	--	0	0	1	13
(a)	alle	8	100	0	0

M = Störungen der Motorik
V = Vitalitätsstörungen
S = Suizidalität

Table 5: Angaben über statistisch signifikante Resultate in erwarteter Form zwischen HIT-Testmerkmalen (HIT) und DIS-Symptomgruppen. Erläuterungen im Text.

Nur bei Q trat in drei Fällen ein erwartetes Ergebnis ein. Bei keinem anderen HIT-Störungsindex ist ein vermehrtes Auftreten psychopathologischer Symptome nachweisbar. Dagegen unterscheiden sich Kontroll- und Patientengruppe in allen acht Symptomgruppen signifikant (Zeile a).

4. Diskussion

Die Daten zeigten für HIT-Variablen weder ausreichende Differenzen im Gruppenvergleich noch genügende Zusammenhänge zu externen Kriterien. Ansatzweise war Trennschärfe zwischen depressiven und gesunden Personen nur nachweisbar für das Integrationsniveau von Klecksdeutungen (I), die behutsame und disziplinierte Nutzung von Farb- und Schattierungsattributen («1»-Kodierungen bei C und Sh), die Relation zwischen sekundär und primär farbde-terminierten C-Antworten im Quotienten C1/(C2 + C3) sowie für die Häufigkeit von Menschenbewegungen (HM). In einem Netz störungsrelevanter externer Daten ließen sich zentrale Testwerte nicht ausreichend «verankern». Oft waren ungünstige HIT-Werte stärker mit vorteilhaften Außenmerkmalen assoziiert als mit störungsrelevanten Aspekten.

Insgesamt wurden 1027 Hypothesen getestet. Nur in 67 Fällen zeigte sich ein erwartetes statistisch signifikantes Resultat. Das sind etwa 6,5 Prozent. Formdeutungen depressiver Menschen heben sich insgesamt nicht in markanter Weise von Antworten psychisch gesunder Personen ab, der vorliegende Untersuchungsbefund spricht daher gegen die Konstruktvalidität der HIT bei Erfassung von Persönlichkeitsmerkmalen, die für Depressionen typisch sind.

Dieser Befund steht nicht in Widerspruch zu bisherigen Studien über HIT-Formdeutungen depressiver Menschen. So fanden Moseley, Duffey & Sherman (1963), die 82 Patienten mit depressiven und schizoaffektiven Störungen testeten, zwischen HIT-Variablen einerseits und selbst- oder fremdbeurteilten depressiven Symptomen und Merkmalen andererseits keinen Zusammenhang. Ausnahme bildete eine faktoranalytisch ermittelte Ähnlichkeit zwischen der Neigung zu Selbstbeschuldigungen (MMPI-Skala «guilt») und den Variablen FA, V und At. Diese Merkmalsgruppe wurde von Moseley et

al. als Ausdruck einer gestörten Beschäftigung mit dem eigenen Körper angesehen, es blieb aber unbegründet, *warum* FA und V Körperschemastörungen anzeigen sollen. Darüber hinaus fanden sich für HIT-Variablen Zusammenhänge nur zu nicht-depressionstypischen externen Merkmalen wie Orientierungsstörungen, Größenideen und paranoid-halluzinatorischen Symptomen. Ein derartiger Befund ist durch die Berücksichtigung schizoaffektiver Psychosen erklärbar, für eine Analyse von HIT-Werten bei depressiver Symptomatik im eigentlichen Sinne aber nicht relevant.

In einer Studie von Endicott und Jortner (1966) hingen C-Antworten mit gutem Formniveau insignifikant positiv, Farbantworten mit schlechtem Formniveau aber signifikant negativ mit Depressivität zusammen. Die Richtung der Zusammenhänge entspricht nicht den Erwartungen, da schlecht strukturierte Farbantworten mit ansteigender Depressivität zunehmen müßten, während «gute» Farbantworten um so häufiger gegeben werden sollten, je schwächer eine Depression ausgeprägt ist. Zudem ist der Befund nur eingeschränkt zur Überprüfung der HIT-Validität bedeutsam, da die Autoren Farbantworten gemäß traditioneller Rorschach-Kriterien signiert hatten. Bei einer zusätzlichen Berechnung nach den Holtzman-Signierungsanleitungen für C ergaben sich keine signifikanten Zusammenhänge zur eingeschätzten Stärke der Depressivität.

Zuckerman, Persky, Eckman & Hopkins (1967) setzten zur Untersuchung depressiver Klinikpatienten vergleichend eine Vielzahl unterschiedlicher Instrumente ein (u. a. Fragebögen, Checklisten), ermittelten für die HIT bei zwölf Analysen nur dreimal signifikante Zusammenhänge zwischen erhöhten Farbscores und anderen Indizes für Depressionen. Für die HIT-Variablen Sh, R (Versager) und RT (Reaktionszeit) wurde kein systematischer Zusammenhang zu Depressivitäts-Maßen gefunden. Andere Instrumente zeigten dagegen ein deutlich höheres Maß an diagnostischer Konkordanz. Dies veranlaßte die Autoren zu der Einschätzung, projektive Testwerte «show a significant relationship to other measures of the trait, but the magnitude of their validity coefficients cannot justify their use in any practical situation». (Zuckerman et al., S. 46).

Die Testautoren (Holtzman et al., 1961) fanden bei der Normierung des Verfahrens an einer depressiven Gruppe (N = 90) verminderte Werte für die Variablen L, FA, C, Sh und M, während V-Scores erhöht waren. Für FD und I erreichten Depressive sogar leicht höhere Werte als Kontrollgruppen. Mit den hier entwickelten theoretischen Vorstellungen sind nur die Veränderungen in FA und V zu erklären. Für L, C, Sh und M können sinnvolle Prognosen nur für Subkategorien der betreffenden Variablen gestellt werden. Beispielsweise ist eine Prognose für den globalen C-Score problematisch, da für C1-Antworten ein Rückgang und für C2- und C3-Antworten ein Anstieg erwartet wird, der relative Anteil beider Tendenzen aber unbekannt ist.

Insgesamt bilden Befunde, die für ein valides, konsistentes Ansprechen von HIT-Variablen auf Persönlichkeitsveränderungen bei Depressionen sprechen, in der Literatur die Ausnahme. Vor diesem Hintergrund und im Lichte der hier vorgelegten Daten muß die Gültigkeit und damit der Wert der HIT in der Diagnostik von Depressionen und Neurosen pessimistisch beurteilt werden. Möglicherweise ließen sich die hier entwickelten diagnostischen Hypothesen bei Personen mit schweren Ich-Defekten (etwa bei schizophrenen oder organisch bedingten Psychosen oder bei geistiger Retardierung) bestätigen. Doch eine fehlende Differenzierung im Bereich normalpsychologischer oder neurotischer Prozesse läßt den Wert eines Verfahrens, das ein subtiles Instrument der Persönlichkeitsdiagnostik zu sein beansprucht, fraglich erscheinen. Nicht auszuschließen ist eine bessere Brauchbarkeit der HIT in inhaltsanalytischen Ansätzen, etwa zur Erfassung verdeckter Ängste oder latenter Aggressivität. Gegenüber der in der Rorschach-Tradition dominierenden formalen Deutungsanalyse bekäme die Formdeutediagnostik bei einem derartigen Interpretationsansatz aber ein völlig neues Profil.

Literatur

- Arieti, S. & Bemporad, J. (1978). *Severe and Mild Depression*. New York: Basic Books. Deutsch. (1983) *Depression*. Stuttgart: Klett-Cotta.
- Anastasi, A. (1982). *Psychological Testing* (5. Auflage). New York: Macmillan.
- Beck, A. T. (1972). *Depression*. Philadelphia: University of Pennsylvania Press.

- Beck, S. J. (1952). *Rorschach's Test*. Vol. III: Advances in Interpretation. New York: Grune and Stratton.
- Beckmann, D. & Richter, H.-E. (1972). *Gießen-Test (GT)*, *Handbuch* (2. Auflage 1975). Bern: Huber.
- Bellak, L., Hurvich, M. & Gediman, H. K. (1973). *Ego Functions in Schizophrenics, Neurotics, and Normals*. London: Wiley.
- Bohm, E. (1957). *Lehrbuch der Rorschach-Psychodiagnostik*. Bern: Huber. Engl. (1958) *Rorschach Test Diagnosis*. New York: Grune and Stratton.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Dilling, C. A. & Rabin, A. I. (1967). Temporal experience in depressive states and schizophrenia. *Journal of Consulting Psychology*, 31, 604–608.
- Emerson, J. D. & Strenio, J. (1983). Boxplots and batch comparison. In D. C. Haaglin, F. Mosteller & J. W. Tukey (Ed.), *Understanding robust and exploratory Data Analysis* (S. 58–96). New York: Wiley.
- Endicott, N. A. & Jortner, S. (1966). Objective measures of depression. *Archives of General Psychiatry*, 15, 249–256.
- Engel, K. (1980). *Zur theoretischen Bedeutung der Variablen der Holtzman-Inkblot-Technik*. Freiburg: Hochschulverlag, 1980.
- Hartmann, H. A. & Rosenstiel, L. v. (Hg.). (1977). *Lehrbuch der Holtzman-Inkblot-Technik (HIT) I* (H. A. Hartmann: Formdeutungsverfahren und Signierungssystem; P. Neumann: Lehrprogramm zur Durchführung, Signierung und Verrechnung), Bern: Huber. Teil II (G. Schmeiser: Dokumentation der Literatur mit Inhaltserschließungen 1956–1977, Bern: Huber, 1980; E. F. Hill: Klinische Interpretation, Bern: Huber 1981).
- Hill, E. F. (1972). *The Holtzman Inkblot Technique. A Handbook for Clinical Application*. San Francisco: Jossey-Bass. Deutsch. (1981) *Klinische Interpretation*. In: H. A. Hartmann & L. v. Rosenstiel (Hrsg.), *Lehrbuch der Holtzman-Inkblot-Technik (HIT) II*. Bern: Huber.
- Hörmann, H. (1964). Theoretische Grundlagen der projektiven Tests. In: R. Heiss (Hg.), *Handbuch der Psychologie* (Bd. 6: Psychologische Diagnostik) (S. 74–112). Göttingen: Hogrefe.
- Hörmann, H. (1982). Theoretische Grundlagen der projektiven Verfahren. In K.-J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie*. (Bd. II) *Psychologische Diagnostik*, (Bd. 1) (S. 173–247). Göttingen: Hogrefe.
- Holtzman, W. H., Thorpe, J. S., Swartz, J. D. & Heron, E. W. (1961). *Inkblot Perception and Personality*, Austin: University of Texas Press, 1961.
- Holtzman, W. H. & Swartz, J. D. (1983). The Holtzman Inkblot Technique: A review of 25 years of research. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 4, 241–259.
- Jäger, R., Lischer, S., Münster, B. & Ritz, B. (1976). *Biographisches Inventar zur Diagnose von Verhaltensstörungen (BIV)*, *Handanweisung*. Göttingen: Hogrefe.
- Jensen, A. R. (1964). The Rorschach technique: A re-evaluation. *Acta Psychologica*, 22, 60–77.
- Klopfer, B., Ainsworth, M. D., Klopfer, W. G. & Holt, R. R. (1954). *Developments in the Rorschach Technique* (Vol. I). New York: Harcourt Brace Jovanovich.
- Koehler, K. & Saß, H. (Hg.). (1984). *Diagnostisches und Statistisches Manual Psychischer Störungen DSM-III*. Weinheim, Beltz.
- Lienert, G. A. (1969). *Testaufbau und Testanalyse*. Weinheim: Beltz.
- Miller, W. R. (1975). Psychological deficit in depression. *Psychological Bulletin*, 82, 238–260.
- Moseley, E. C., Duffey, R. F. & Scherman, L. J. (1963). An extension of the construct validity of the Holtzman Inkblot Technique. *Journal of Clinical Psychology*, 19, 186–192.
- Neumann, P. (1977). Beiheft zur Durchführung, Signierung und Verrechnung. In H. A. Hartmann & L. v. Rosenstiel (Hrsg.), *Lehrbuch der Holtzman-Inkblot-Technik (HIT) I*, Bern: Huber.
- Peterson, R. A.: (1978). Holtzman Inkblot Technique. In O. K. Buros (Ed.), *The Eighth Mental Measurement Yearbook*. Highland Park: The Gryphon Press.
- Robins, L. N., Helzer, J. E. Croughan, I. & Ratcliff, K. S. (1981). The NIMH Diagnostic Interview Schedule: It's history, characteristics, and validity. *Archives of General Psychiatry*, 38, 381–389.
- Salzman, L., Goldstein, R., Atkins, R. & Babigan, H. (1966). Conceptual thinking in psychiatric patients. *Archives of General Psychiatry*, 14, 55–59.
- Schmeiser, G. (1985). *Bewährungsfelder eines Formdeutungsverfahrens – Eine Bestandsaufnahme zur Gültigkeit der Holtzman-Technik*. Dissertation, München.
- Siegel, S. (1976). *Nicht-parametrische statistische Methoden*. Frankfurt am Main: Fachbuchhandlung für Psychologie.
- Spitznagel, A. (1982). Grundlagen, Ergebnisse und Probleme der Formdeutungsverfahren. In K.-J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie*, (Bd. II), *Psychologische Diagnostik*, (Bd. 3) (S. 186–257). Göttingen: Hogrefe.
- Suinn, R. M. & Oskamp, S. (1969). *The predictive validity of projective measures: a fifteen-year evaluative review of research*. Springfield: Thomas.
- Tölle, R. (1982). *Psychiatrie*. (6., neuverfaßte und erweiterte Auflage). Berlin: Springer.
- Witichen, H.-U. & Rupp, H. U. (1981). *Das Diagnostic Interview Schedule (Version II), Deutsche Version*. München: Max-Planck-Institut für Psychiatrie.
- Zubin, J. (1954). Failures of the Rorschach technique. *Journal of Projective Techniques*, 18, 303–315.
- Zuckerman, M., Persky, H. Eckman, K. M. & Hopkins, T. R. (1967). A multitrait multimethod measurement approach to the traits (or states) of anxiety, depression and hostility. *Journal of Projective Techniques and Personality Assessment*, 37, 39–48.