

Therapieevaluation **am Dan Casriel Institut**

Therapieergebnisse von fünf NIP Workshops im **Jahr 1995**

Auswertung und Ergebnisdarstellung der Evaluation:

Forschungsstelle an den Psychosomatischen Kliniken Grönenbach,

Dipl.-Psych. Robert Mestel

Mai bis Oktober 1996

Unveröffentlichter Bericht; Grönenbach, 16.10.96

Zusammenfassung:

Am Dan Casriel Institut, geleitet von Dr. med. Ingo Gerstenberg, wurden fünf Psychotherapie-Workshops des Jahres 1995 hinsichtlich der Veränderungen der Teilnehmer beurteilt. Dazu wurden den Gruppenteilnehmern zu Beginn der Workshops und etwa einen Monat danach Fragebogen zur psychischen Symptomatik (Symptom Checkliste: SCL-90-R) und zu interpersonellen (zwischenmenschlichen) Problemen (Inventar zur Erfassung interpersoneller Probleme: IIP) vergeben. Die Meßergebnisse beider Zeitpunkte wurden mittels vier verschiedener statistischer Methoden miteinander verglichen und es zeigte sich, daß die allermeisten der 117 Personen zum Nachbefragungszeitpunkt weniger unter psychischen Symptomen, wie z.B. Depressionen, Ängsten, sozialer Unsicherheit oder Zwanghaftigkeit litten. 50-73% der Personen waren zum Nachbefragungszeitpunkt im Symptombereich klinisch relevant gebessert, d.h. sie waren dann in bezug auf die Höhe der Symptomskalen nicht mehr von einer gesunden Vergleichsstichprobe zu unterscheiden. 72% der Patienten zeigten statistisch signifikant weniger generelle Symptombeschwerden über den Zeitverlauf. Im Bereich interpersoneller Probleme ergaben sich insgesamt und in Einklang mit der relevanten Vergleichsliteratur weniger Veränderungen. Zum Nachbefragungszeitpunkt zeigten sich jedoch 68-85% der Teilnehmer als klinisch relevant gebessert. Einerseits gaben die Teilnehmer an, weniger an Problemen mit sozialem Rückzug („Introversion“) zu leiden. Jedoch ergab sich auch, daß einige Personen zum Nachbefragungszeitpunkt mehr Probleme mit „Dominanz“, „Streitsüchtigkeit“ und „Expressivität“ aufwiesen wie vorher. In einer statistischen Detailanalyse konnte jedoch gezeigt werden, daß nur 40% der im zwischenmenschlichen Bereich verschlechterten Teilnehmer *gleichzeitig* auch im Symptombereich schlechter bzw. gleichgeblieben waren. Dieses Ergebnis könnte dahingehend interpretiert werden, daß diese Teilnehmer die genannten Probleme im interpersonellen Bereich nun stärker wahrnehmen und empfinden oder, daß sie (vorläufig) über das Therapieziel (z.B. sich abgrenzen lernen) „hinausgeschossen“ sind und nun vermehrt mit ihrer sozialen Umgebung in Konflikte geraten.

Insgesamt decken sich die Ergebnisse mit denen von anderen Therapiestudien: Der Symptombereich ändert sich schneller und in kürzerer Zeit stärker im Vergleich zum zwischenmenschlichen Bereich. Änderungen auf interpersonellen Skalen sind schwer pauschal zu beurteilen und sollten individueller und mit zusätzlichen anderen Instrumenten neben Fragebogen erfaßt werden.

Mit der vorliegenden Pilotstudie liegt somit die erste gruppenstatistische Untersuchung zur Therapiemethode „New Identity Process“ vor.

1. Einleitung und Ausgangslage

Die meisten der bisherigen Therapieschulen entstanden in der Mitte dieses Jahrhunderts und wurden von „charismatischen Persönlichkeiten“ begründet, die keinen Wert auf die wissenschaftliche Überprüfung ihrer Annahmen und Veränderungsmodelle in Psychotherapien legten. Grawe (1994) spricht hier von „psychotherapeutischen Theorien erster Generation“. Die Überprüfung der Effektivität der Therapien war bzw. ist für die meisten Therapieschulen außer der Verhaltenstherapie, der Gesprächspsychotherapie und zum Teil der Psychoanalyse dadurch erschwert, daß sie in erster Linie nicht im universitären Umfeld entwickelt wurden. Auch für die Therapieform „New-Identity-Process“ (NIP; bzw. „Casriel“- oder „Bonding“ Therapie) liegt nach dem heftig diskutierten Buch von Grawe, Bernati & Dornauer (1994, S. 735) keine einzige bisherige stichhaltige Wirksamkeitsuntersuchung und damit das Minimal Kriterium vor, daß man von einer wissenschaftlich fundierten Therapieform sprechen kann. Außer den Einzelfallberichten durch den Begründer Casriel selbst (Casriel, 1975) oder den New Yorker Psychotherapeuten Holtzman liegen keine veröffentlichten Gruppenuntersuchungen oder gar kontrollierte Studien vor. Eine Untersuchung zur Veränderung der Bindungsstile (im Sinne von John Bowlby) bei Patienten in stationärer psychosomatischer Behandlung im Laufe einer speziellen NIP Gruppe von Markert & Mestel (in Vorb.) wird gerade fertiggestellt. Im stationären Setting besteht jedoch das grundlegende Problem, daß neben der Spezialgruppe „Bonding“ (NIP) noch zahlreiche andere Therapieangebote vorliegen und natürlich auch wirken. Zwar wird NIP nach den Befunden von Wittmann, Held, Rudolf & Schulze (1996, S. 77) von den Patienten als ein sehr hilfreicher Wirkfaktor im Vergleich zu anderen Angeboten angesehen, aber es war in dieser Studie *nicht* möglich, mit der „Dosis“ (Anzahl) an erhaltenen NIP Sitzungen in der stationären Therapie den Therapieerfolg vorherzusagen.

Der häufige Einsatz der Methoden des NIP in vielen Ländern und in unterschiedlichsten Anwendungsgebieten (z.B. Suchttherapie, Psychosomatik, Selbsterfahrung) zeigt, daß klinische Praktiker vom Nutzen für ihre Patienten überzeugt sind.

Emotionen sind ein in der Psychotherapieforschung nur spärlich untersuchtes Thema (siehe Greenberg & Safran, 1987). Diese Tatsache könnte damit erklärt werden, daß die "emotional-kathartisch" arbeitenden Therapeuten nicht forschungsorientiert sind oder bisher sogar gegen Forschung eingestellt waren. Betrachtet man die Ergebnisse wissenschaftlicher Studien zur Effektivität vom Ausdruck von Gefühlen oder „Katharsis“ in Psychotherapien, dann zeigen sich in den wenigen Studien höchst widersprüchliche Befunde. Zwar gilt eine hohe affektive Erregung (arousal) als eine gute Vorbedingung für Einstellungsänderungen (Frank, 1973), aber nach Greenberg et al. (1987) oder Yalom (1992, S. 89ff) ist der pure Ausdruck von Emotionen nur ein temporärer "Streßreduzierer" *und erzeugt keine anhaltenden Veränderungen* (S.198).

Als (sehr allgemeine) Zusammenfassung schreiben Greenberg et al. (1987, S.101, Übersetzung durch R.M.):

"Emotionale Erregung und emotionaler Ausdruck hat unter bestimmten Umständen und bei bestimmten Typen von Individuen und Problemen einen Einfluß auf Veränderungen".

Folgende Punkte werden an der vorliegenden Katharsisforschung kritisiert:

1. Konzeptuelle und methodologische Schwächen: Noch nie wurden Kontrollgruppen zum Vergleich herangezogen.
2. Heterogenität der Stichproben.
3. Es bleibt unklar, was im Therapieprozess genau wirkt und wie emotionales Erleben zu Veränderungen in der Psychotherapie führt.
4. Statt zu messen, wie oft oder häufig der affektive Ausdruck passiert bzw. vorkommt, sollte man den exakten Kontext (oder die Bedeutung) des affektiven Ausdrucks beachten. Zu untersuchen wären auch die Auftrittsmuster verschiedener affektiver Prozesse (z.B. Übergang von Wut zu Schmerz).
5. Die bisher entwickelten Instrumente sind unzureichend, um die Intensität und den Wechsel der Emotionen zu messen.
6. Es fehlt eine integrative Rahmentheorie über die Rolle der Emotionen in der Psychotherapie.

NIP Ausbilder und praktizierende NIP Therapeuten betonen jedoch, daß die Methode des New Identity Process wesentlich mehr Elemente umfaßt, als das hervorsteckende, und damit für diese Therapieform charakteristische und einzigartige Merkmal des „emotionalen Ausdrucks bei körperlicher Nähe“, das eben im Zusammenhang mit den „Katharsisbefunden“ andiskutiert wurde. Nach Casriel's Theorie setzt die Therapieform neben dem emotionalen Bereich auch auf der Verhaltensebene („Einüben neuer Verhaltensmuster“) und der kognitiven Ebene („Einstellungsarbeit“) an. NIP kann somit, wie auch Holtzman schreibt, als integrative, oder zumindest eklektische Therapieform verstanden werden.

NIP wird in verschiedenen Settings (z.B. stationär oder ambulant) angewendet und es existiert nach Kenntnis des Autors kein schriftliches, verbindliches und standardisiertes Manual, an das sich alle praktizierenden NIP-Therapeuten halten.

2. Ziel der Untersuchung:

Mittels eines naturalistischen Designs sollte festgestellt werden, ob sich Teilnehmer an NIP-Kursen am Dan Casriel Institut über einen Zeitraum von 1-2 Monaten nach dem Workshop auf der Ebene der Symptomatik und im Bereich von Beziehungsproblemen verändern. Ferner wurde untersucht, in welchen spezifischen Bereichen sich die Veränderungen ergeben und wie stark sie sind.

Bereits hier müssen kurz die wichtigsten Einschränkungen der Aussagekraft dieser Untersuchung erwähnt werden (näheres im Kapitel „Einschränkungen“ am Ende):

Es kann hier *nicht* bewiesen werden, welchen *spezifischen* Einfluß die Methode NIP oder bestimmte Interventionstechniken auf die Veränderungen gehabt haben, da keine unbehandelte Kontrollgruppe untersucht wurde und auch der Einfluß von weiteren ambulanten Therapien neben NIP auf die Ergebnisse vor und nach dem Workshop nicht abgeschätzt werden konnte.

Diese Arbeit soll somit als eine erste Pilotstudie mit eher explorativem Charakter aufgefaßt werden und kann somit als Grundlage für die Erforschung spezifischerer Annahmen und Hypothesen zur Wirkungsweise der NIP Behandlungsmethode herangezogen werden.

3. Methoden und Durchführung der Untersuchung

An der Untersuchung nahmen insgesamt 136 Teilnehmer von fünf Workshops teil:

Workshop A (21.-29.7.95; 33 Personen),
 B (5.-9.9.95; 27 Personen),
 C (26.-29.9.95; 20 Personen),
 D (17.-22.11.95; 30 Personen),
 E (17.-22.11.95 und 17.-27.11.95; 26 Personen).

Von den 136 Teilnehmern waren 82 (60,3%) Frauen und das durchschnittliche Alter betrug 39,8 Jahre (Standardabweichung: 9.1). 32% hatten Real-oder Hauptschule, 21% Abitur und 47% Hochschulabschluß.

Die Teilnehmer der fünf Workshops unterschieden sich nicht in bezug auf das Alter oder den Bildungsgrad, jedoch waren mehr Frauen in den Workshops A und B und mehr Männer in der Gruppe E.

Allen 136 Personen wurden zu Beginn des Workshops der Fragebogen zu körperlichen und psychischen Symptomen „SCL-90-R“ (Symptom-Checkliste 90-revidiert; Franke, 1995; 90 Fragen mit fünfstufigen Antwortskalen; Skalen z.B. Depressivität, Unsicherheit im Sozialkontakt, Somatisierung) und das „Inventar Interpersoneller Probleme“ („IIP“; Horowitz, Strauß & Kordy, 1994; 64 Fragen mit fünf stufigen Antwortskalen; Skalen z.B. Probleme mit Unterwürfigkeit, Streitsüchtigkeit, Fürsorglichkeit) vorgelegt.

Über die Art und die Zahl der Personen, die an den Workshops teilnahmen, die Fragebogen jedoch bereits zu Beginn *nicht* ausfüllten, können keine Angaben gemacht werden.

Aus den Angaben oben wird ersichtlich, daß die NIP Workshops ca. eine Woche andauerten. Die Inhalte der Gruppen (Thematik, Interventionen etc.) können den Programmen des Dan Casriel Instituts (1995) entnommen werden. Nähere Einzelheiten zur Art der Behandlungsform sind von den Dozenten oder dem Leiter der Einrichtung Dr. Ingo Gerstenberg zu erhalten. Die Workshops wurden nach Kenntnis des Autors ohne manualisierte Behandlungspläne durchgeführt und es gab verschiedene Gruppenleiter. Inwieweit also „wirklich“ NIP Therapie durchgeführt wurde, kann nicht genau festgestellt werden.

Die Teilnehmer wurden vom Gruppenleiter ausdrücklich um die Teilnahme an der Untersuchung gebeten und auf das nachfolgende Anschreiben vorbereitet.

Den 136 Personen wurden die zwei Fragebogen nach 1-2 Monaten erneut postalisch verschickt. Es antworteten bei nur einer Befragung 117 Personen, was einer sehr hohen Rücklaufquote von 87.3% entspricht. Die Generalisierung der im folgenden zu berichtenden Ergebnisse für die Antworter auf die gesamte Ausgangsgruppe ist somit prinzipiell möglich, vor allem, da sich die 19 Nicht-Antworter von den Antwortern nur in einer von 18 Skalen statistisch signifikant unterschieden: Die Nicht-Antworter hatten bei Beginn der Workshops eher mehr Probleme damit, Anderen Gefühle zu zeigen und Anderen gegenüber zu kalt zu sein (Skala „abweisend“ im

IIP). Auch auf den sozio-demographischen Variablen gibt es keine Unterschiede zwischen Antwortern und Nicht-Antwortern.

Die Fragebogen wurden vom Sekretariat des Instituts gesammelt und dann zur Eingabe und Skalenverrechnung an ein externes unabhängiges Institut, die Forschungsstelle an den Kliniken für Psychosomatische Medizin Grönenbach weitergeleitet. Von der wissenschaftlichen Hilfskraft Ursel Partsch wurden die *einzelnen* Ergebnisse *aller* Patienten auf übersichtlichen Grafiken dargestellt. Dabei wurde auch markiert, in welchen einzelnen Bereichen sich die Patienten verbessert oder verschlechtert hatten. So war es den interessierten Mitarbeitern am Dan Casriel Institut möglich, sich von allen Einzelpersonen ein Bild über die weitere Entwicklung in verschiedenen Lebensbereichen zu machen.

Zusätzlich wertete Dipl.-Psych. Robert Mestel die Daten gruppenstatistisch anhand von vier Methoden aus:

1. Signifikanzstatistik (Varianzanalysen mit Meßwiederholung: Prüfung der Unterschiede der Mittelwerte von Prä und Post).
2. Effektstärkenberechnung (Standardisierte Mittelwertsdifferenzen).
3. Prozentsatz einzelner Personen, die sich statistisch signifikant verändert haben.
4. Prozentsatz einzelner Personen, die sich klinisch relevant verändert haben.

Die Berechnungen erfolgten mit dem statistischen Programmpaket SAS for Windows, Version 6.0.

4. Ergebnisse

Vorweg wurde geprüft, ob es Unterschiede im Therapieerfolg zwischen den fünf Workshops gab. Es fanden sich auf keiner der 18 Skalen entsprechende Unterschiede: Alle fünf Gruppen profitierten in ähnlicher Weise.

4.1. Bereich der psychischen Symptome

4.1.1. Mittelwertsunterschiede auf den Symptomskalen zwischen Prä (Workshopbeginn) und Post (Nachbefragung)

Die SCL-90-R Skalenantworten gehen jeweils von 0 (Symptom nicht vorhanden) bis 4 (sehr stark vorhanden). Alle 10 Skalen sind so definiert, daß sie Problembereiche erfassen (z.B. Skala 6: *Probleme mit Aggressivität und Feindseligkeit*). Die Skalenbezeichnungen sind zumeist anschauliche Oberbegriffe der Einzelitems (hohe inhaltliche Validität). Unter „Psychotizismus“ versteht man eine kontinuierliche Dimension des menschlichen Erlebens, die von verzerrtem, isoliertem, schizoidem Lebensstil bis hin zu Primärsymptomen der Schizophrenie wie Halluzinationen und Gedankenzerfall reicht (Franke, 1995, S.23). Die generelle Symptomatik besteht aus dem Mittelwert aller neun anderen Symptomskalen.

Wie aus Abbildung 1 ersichtlich, zeigten sich für alle Teilnehmer im Durchschnitt auf allen 10 Symptomskalen starke Verbesserungen von Prä nach Post. Ferner zeigt sich, daß die Teilnehmer zum Post Zeitpunkt in sechs Skalen im Normbereich der gesunden Normstichprobe (n=1002) liegen und sich in den übrigen Skalen deutlich auf den Normbereich zubewegt haben.

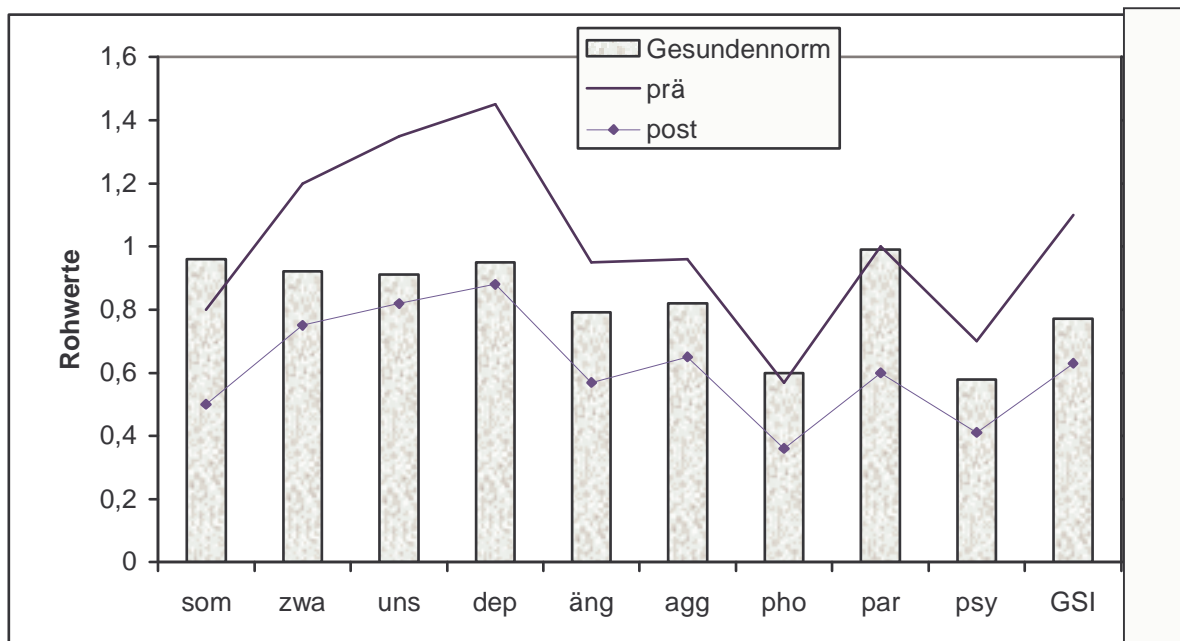


Abb. 1: Prä-Post Vergleich auf den Symptomskalen der SCL-90-R.

Anmerkungen: n=117; Gesundennorm: Innerhalb dieser Fläche liegen 68% der Normalpersonen. Som=Somatisierung, zwa=Zwanghaftigkeit, uns=Unsicherheit im Sozialkontakt, dep=Depressivität, äng=Ängstlichkeit, agg=Aggressivität/Feindseligkeit, pho=Phobische Angst, par=Paranoides Denken, psy=Psychotizismus, GSI=Globaler Schwere Index

4.1.2. Effektstärkenberechnungen für die Symptomskalen

Effektstärken drücken etwas über die Größe des Unterschieds zwischen zwei Populationen aus. Sie sind im Gegensatz zu Signifikanztests (Abb. 1) unabhängig von der Stichprobengröße und werden häufig herangezogen, um Therapiestudien miteinander vergleichen zu können. Der Effektstärkenberechnung wurde die von W.W. Wittmann (1996) empfohlene, und bereits z.B. durch Tscheulin & Schäfer (1993) erprobte Formel „Differenz zwischen Prä und Katamnesewert, geteilt durch die Standardabweichung der Differenzen“, zugrunde gelegt. Diese Berechnungsweise kommt zu nahezu identischen Ergebnissen wie die von Hartmann und Herzog (1995) empfohlenen Berechnungen von Mac Gaw & Glass (1980) und Hartmann, Herzog & Drinkmann (1992) (siehe Mestel, 1996). Bei der Interpretation der Höhen der Effektstärken ist jedoch darauf zu achten, daß es sich um Prä-Katamnese Vergleich derselben Stichprobe handelt, also keine Kontrollgruppe zur Effektstärkenberechnung herangezogen wurde. Dadurch können die Cohenschen Interpretationsrichtlinien nur mit Vorbehalt angewendet werden. Es mehren sich die Hinweise, daß die Effektstärkenberechnungen mit versus ohne Kontrollgruppen völlig unterschiedliche Ergebnisse erzeugen (mit Kontrollgruppe: siehe Wittmann & Matt, 1986; ohne Kontrollgruppe siehe Hauzinger, de Jong-Meyer, Treiber, Rudolf & Thien, 1996). Bis es Meta-Analysen von Therapiestudien gibt, die auf Prä-Post bzw. Prä-Katamnese Effektstärkenberechnungen basieren, und die neue Interpretationsstandards setzen wird von uns und anderen vorgeschlagen, von den Prä-Post

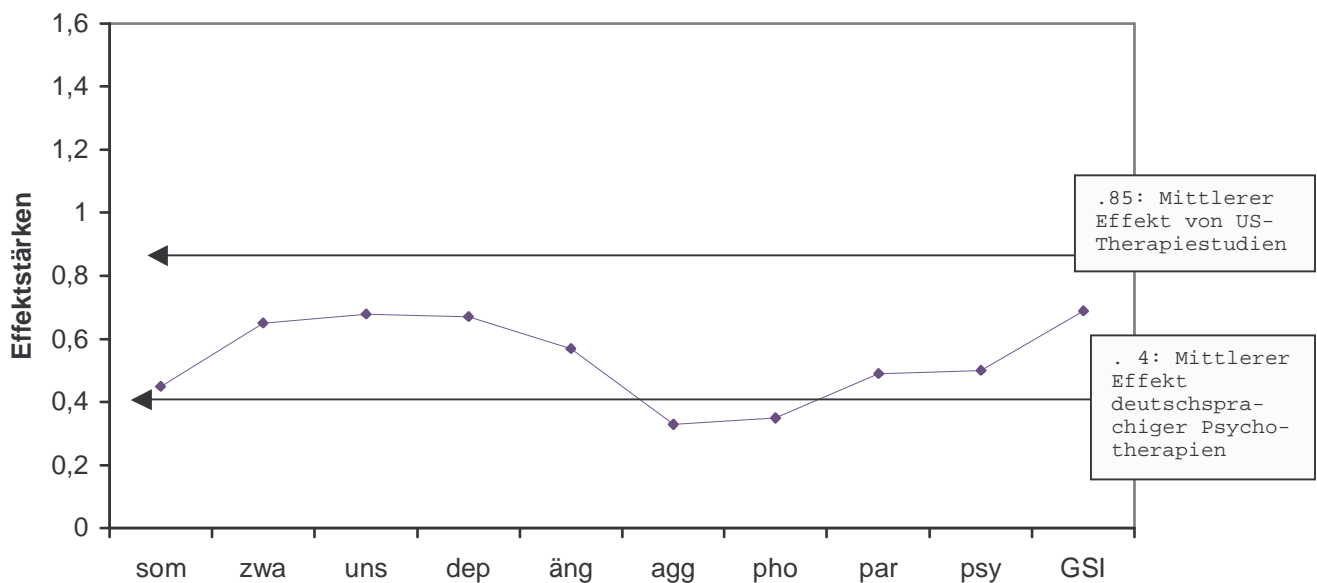


Abb. 2: Effektstärkenberechnungen für die Symptomskalen (Standardisierte prä-post-Differenzen).

Anmerkungen: n=117; Som= Somatisierung, zwa= Zwanghaftigkeit, uns= Unsicherheit im Sozialkontakt, dep= Depressivität, äng= Ängstlichkeit, agg= Aggressivität/ Feindseligkeit, pho= Phobische Angst, par= Paranoides Denken, psy= Psychotizismus, GSI= Globaler Schwere Index; Effektstärkeeinteilung: Kleiner Effekt: 0,2-0,5; mittlerer Effekt: 0,51-0,8; großer Effekt >.8.

Effektstärken die Effektstärken von unbehandelten Kontrollgruppen abzuziehen, um sie dann anhand von Cohen's Richtlinien (ersichtlich aus den Abbildungen 2 und 5) bewerten zu können. Diese durchschnittlichen Kontrollgruppeneffektstärken werden mit $ES=.123$ von Hartmann et al. (1996, S. 339) veranschlagt und liegen bei Grawe (1992) in ähnlicher Höhe.

Die Effektstärkendarstellung (Abbildung 2) zeigt, daß auf den Skalen „Zwanghaftigkeit“, „Unsicherheit im Sozialkontakt“, „Depressivität“, „Ängstlichkeit“ und „generelle Symptomatik“ mittlere Effektstärken erreicht wurden und auf den übrigen Skalen kleine Effekte. Die Effektstärkenhöhen decken sich etwa mit den Durchschnittswerten deutschsprachiger Psychotherapiestudien (Wittmann & Matt, 1986).

4.1.3. Klinisch relevante Veränderungen (Symptomskalen)

In der folgenden Tabelle 1 wird der Anteil der Teilnehmer (in Prozent) genannt, die zum Nachbefragungszeitpunkt im Normbereich (68% der gesunden Normstichprobe; $n=1002$) lagen, sich somit nicht mehr von den „Gesunden“ unterschieden. Personen, die am Ende bzw. nach einer Intervention (z.B. Therapie) dieses Kriterium erfüllen gelten nach Jacobson, Follette & Revensdorf (1984) als klinisch relevant gebessert.

Tabelle 1: Prozentsatz klinisch relevant gebesserter Personen zum Nachbefragungstermin (Symptomskalen)

SCL-90-R Skala	Prozentsatz klinisch relevant gebesserter Personen von $n=117$
1. Somatisierung	72.6%
2. Zwanghaftigkeit	71.8%
3. Unsicherheit im Sozialkontakt	59.8%
4. Depressivität	58.1%
5. Ängstlichkeit	66.7%
6. Aggressivität und Feindseligkeit	50.4%
7. Phobische Angst	70.1%
8. Paranoides Denken	69.2%
9. Psychotizismus	63.2%
10. Generelle Symptomatik	57.3%

Es zeigt sich, daß die überwiegende Zahl der Teilnehmer im Symptombereich zum Nachbefragungstermin klinisch relevante Verbesserungen aufweisen. Es muß jedoch bedacht werden, daß einige Gruppenteilnehmer vor allem zu Selbsterfahrungszwecken an den Workshops teilnahmen. Diese oder auch andere Teilnehmer waren bereits zu Beginn der Workshops psychopathologisch unauffällig im Sinne von nicht von der Norm abweichenden Symptombelastungsangaben. Deshalb wird in der Folge (Kapitel 4.1.4.) auch die Zahl der einzelnen Personen bestimmt, die *statistisch* signifikante Veränderungen im Symptombereich aufwiesen.

4.1.4. Statistisch signifikante Veränderungen einzelner Personen auf den Symptomskalen.

In Abbildung 3 zeigt sich pro Symptomskala der Prozentsatz der Personen, die statistisch signifikant verbessert, statistisch signifikant verschlechtert oder gleichgeblieben sind. Als gleichbleibende, unveränderte Werte gelten solche Prä-Post Differenzen, die nicht von den Zufallsschwankungen des Meßinstruments (SCL-90-R), welches nur mehr oder weniger genau mißt, unterschieden werden können. Wenn hier von Verbesserung oder Verschlechterung die Rede ist, so wird diesen Aussagen eine Irrtumswahrscheinlichkeit von 5% zugrunde gelegt (Berechnungen nach den Angaben zur kritischen Differenz im SCL Testmanual, S.28). Die meisten statistisch signifikanten Verbesserungen (72%) gibt es auf der Skala „Generelle Symptomatik“, gefolgt von „Depressivität“ (58%), „Unsicherheit im Sozialkontakt“ (48%) und „Zwanghaftigkeit“ (45%). Andererseits zeigt sich auch, daß auf den Skalen 1, 6, 7 und 8 etwa 70% der Personen zumindest auf der Symptomebene unverändert bleiben.

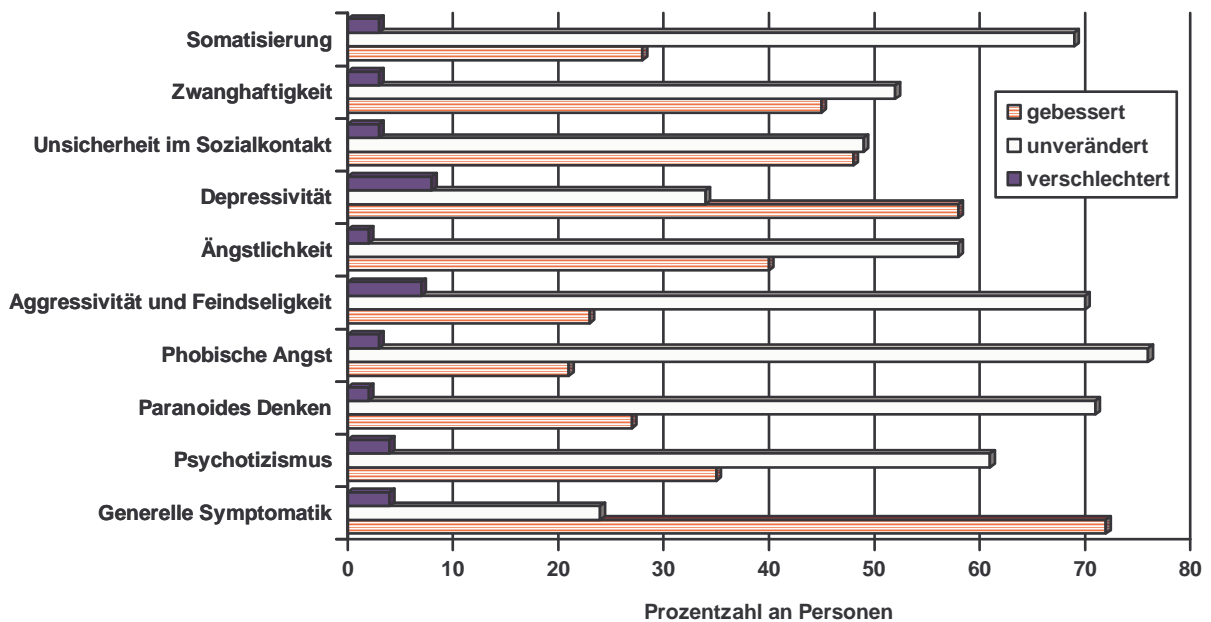


Abbildung 3: Prozentsätze statistisch signifikant veränderter Einzelpersonen auf den Symptomskalen.

4.2. Bereich interpersoneller Probleme

4.2.1. Prä-Post Mittelwertsvergleiche und Effektstärkenberechnungen für die interpersonellen Skalen

Auf vier von acht Skalen des Inventars interpersoneller Probleme (IIP) gab es statistisch signifikante Unterschiede zwischen den Prä und Post Gruppenmittelwerten. Im IIP sind die Skalen explizit als „Probleme mit ...“ (z.B. zuviel Streitsüchtigkeit) definiert, so daß hohe Stanine Normwerte mehr Probleme bedeuten und niedrige entsprechend weniger Probleme mit diesem Bereich. Wie aus Abbildung 4 ersichtlich, verbessern sich die Teilnehmer durchschnittlich im Bereich „Probleme mit Introvertiertheit“ (Rückzug), verschlechtern sich jedoch in den drei Kontroll-Skalen (Im „Norden“ des Zirkumplex Modells“) „PA, BC und NO“. Insgesamt jedoch bewegen sich alle Prä- oder Post MW's zwischen 4,5 und 5,5 Stanine Normwerten, liegen also völlig im Durchschnitt der Normpopulation. Es zeigt sich, daß die interpersonellen Probleme sich wesentlich langsamer und weniger verändert haben im Vergleich zu den Symptomskalen. Vergleiche zwischen den Effektstärkenabbildungen 2 und 5 verdeutlichen dies. Dieser Befund deckt sich mit den Ergebnissen der großen deutschen multizentrischen Studie zu interpersonellen Problemen (Keller & Schneider, 1993): Auch hier ergab sich, daß selbst durch intensive stationäre Psychotherapie interpersonelle Probleme nur mäßig zu beeinflussen waren.

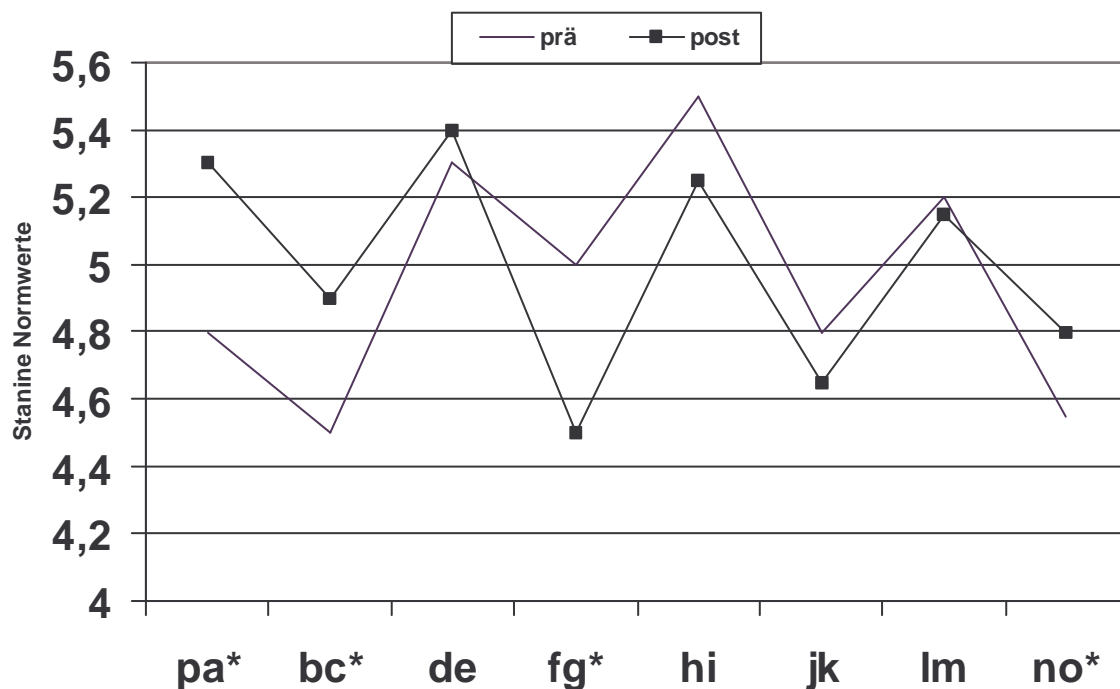


Abbildung 4: Prä-Post Vergleich mit den interpersonalen Skalen (n=117).

Anmerkungen: Mittelwert in der Normbevölkerungsstichprobe: 5; Streuung: 1,96; pa=Autokratisch, bc=streitsüchtig, de=abweisend, fg=introvertiert, hi=unterwürfig, jk=ausnutzbar, lm=fürsorglich, no=expressiv.

*: Statistisch signifikanter Unterschied zwischen den prä und post-Werten ($p < .05$).

Durch die Arbeitsgruppe von Ken Howard (siehe Lueger, 1995) ist seit längerem bekannt, daß sich in Psychotherapien zuerst die Befindlichkeit (Wohlbefinden), dann die Symptomatik und dann erst die Funktionsstörungen (z.B. Partnerschaftskonflikte) verändern. Veränderungen im interpersonellen Bereich sind somit weniger schnell zu erwarten.

Schwer zu interpretieren ist die durchschnittliche signifikante Verschlechterung der Gesamtgruppe auf drei IIP Skalen. Der Testautor Bernhard Strauss selbst (mündliche Mitteilung an R.M. am 8. 6. 96) interpretiert leichte IIP Verschlechterungen auch als „gestiegene interpersonelle Sensitivität“. Die betroffenen Personen würden in den gestiegenen IIP Werten andeuten, daß sie den betreffenden Bereich (z.B. Expressivität) erst jetzt richtig als Problem sehen und benennen könnten. Dieses Argument bringt jedoch nur bedingt weiter, da es ebenso auf die verbesserten Fälle angewendet werden kann: Vielleicht „verdrängen“ die statistisch signifikant verbesserten Personen nur ihre „wahren“ interpersonellen Probleme, die sie jedoch (aus der Fremdperspektive durch andere Personen beurteilt) weiterhin besitzen.

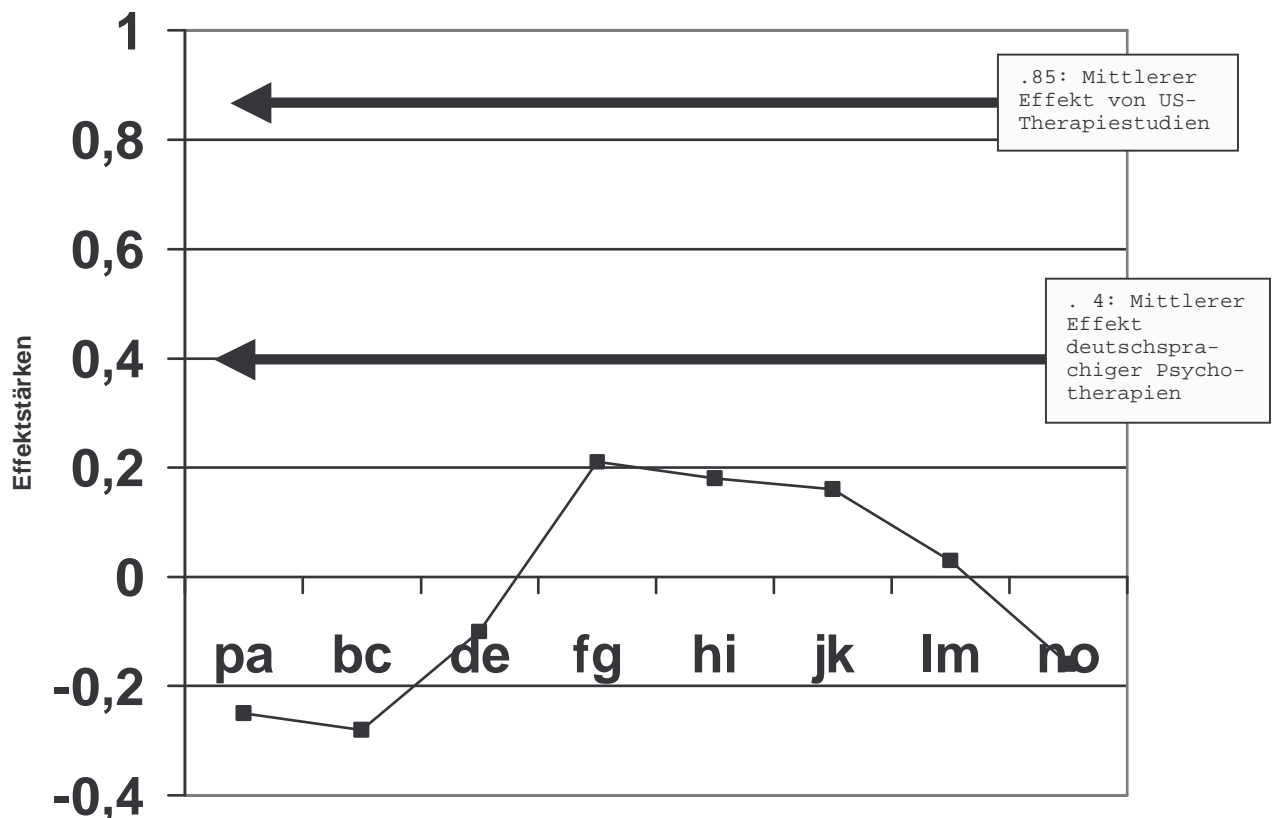


Abbildung 5: Effektstärken für die interpersonellen Skalen (Standardisierte Prä-Post-Differenzen).

Anmerkungen: pa=Autokratisch, bc=streitsüchtig, de=abweisend, fg=introvertiert, hi=unterwürfig, jk=ausnutzbar, lm=fürsorglich, no=expressiv.

Effektstärkeinteilung: Kleiner Effekt: 0,2-0,5; mittlerer Effekt: 0,51-0,8; großer Effekt >.8.

Um die Befundlage besser interpretieren zu können, wurde überprüft, ob die *einzelnen* in den IIP Skalen verschlechterten Personen sich *gleichzeitig* auch in der SCL-90-R verschlechtert haben bzw. gleichgeblieben sind. Von den 34 Personen, die auf allen vier Skalen (PA, BC, DE und NO) verschlechtert waren¹ oder gleiche Werte behielten (Post Wert war ein SN Wert schlechter oder gleich wie Prä Wert), waren 41% (N=14) auch auf der SCL-90-R Skala „generelle Symptomatik“ verschlechtert bzw. gleichbleibend².

Dagegen waren 59% (n=20) der interpersonell eindeutig „Verschlechterten“ in der generellen Symptombelastung deutlich bis sehr deutlich verbessert. Diese Gruppe von 14 Personen, also 12% der Antwortter insgesamt stellt somit die eigentliche „wirkliche“ Problemgruppe dar.

Zusammenfassend läßt sich feststellen, daß es bei ca. 40% aller im zwischenmenschlichen Bereich verschlechterten Personen auch im Symptombereich keine Besserungen bzw. sogar Verschlechterungen gab.

Die Verschlechterungen lagen inhaltlich klar eher auf dem Kontrollpol und der feindseligen Seite. Es bleibt in weiteren Studien festzustellen, ob dies ein spezifischer Charakteristischer Effekt für die NIP Methode oder deren Teilnehmer/Patienten ist, oder ob sich derselbe Effekt auch in anderen Therapieschulen bzw. anderen Patientengruppen finden läßt. Eine mögliche Erklärung für den genannten Befund könnte sein, daß einige Teilnehmer, gerade häufiger Frauen (siehe nächster Absatz), durch die zahlreichen Abgrenzungsübungen z.B. bei der Einstellungsarbeit über das Ziel „hinausgeschossen“ sind und mit ihrer neu gestärkten Selbstbehauptung mehr Probleme mit Streitsüchtigkeit, Dominanz und Expressivität aufweisen.

Interessanterweise ergaben weitere Detailanalysen, daß sich Männer stärker auf der Skala PA (Probleme mit Autokratie und Dominanz) verbesserten im Vergleich zu Frauen. Dies war der einzige Geschlechtsunterschied auf den 18 Skalen hinsichtlich den Veränderungsprozessen.

¹ definiert durch: Prä Wert \geq 2 SN Werte im Vergleich zum Post Wert; orientiert sich an der kritischen Differenz mit $\alpha=10\%$.

² Verschlechterung wurde hier definiert als: Post-Prä Differenzwert > 0.17 Rohwerte (kritische Differenz bei $\alpha=5\%$; siehe Testmanual).

Als gleichbleibend galten Post-Prä Differenzen ≥ 0 und < 0.17 .

4.2.2. Klinisch relevante Veränderungen auf den interpersonellen Problembereichen

Analog zum Kapitel 4.1.3. wird in Tabelle 2 der Prozentsatz klinisch relevant gebesserter Patienten im interpersonalen Bereich angegeben: Personen mit einem Post Wert von SN < 7 gelten als nicht mehr von der Gesunden Stichprobe (Mittelwert=5; Standardabweichung=1.96; d.h. 68% liegen zwischen 3.04 und 6.96) verschieden.

Es zeigt sich, daß zum Post Zeitpunkt nur noch ca. 20-30% der Patienten stark erhöhte Werte in den IIP Skalen aufweisen.

Tabelle 2: Prozentsatz klinisch relevant verbesserter Personen zum Nachbefragungstermin.

Inventar interpersoneller Probleme: Skalen	Prozentsatz klinisch relevant gebesserter Personen von n=117
Autokratisch	68.4%
2. Streitsüchtig	75.2%
3. Abweisend	76.1%
4. Introvertiert	82.1%
5. Unterwürfig	74.4%
6. Ausnutzbar	84.6%
7. Fürsorglich	73.5%
8. Expressiv	80.3%

4.2.3. Statistisch signifikante Veränderungen auf den interpersonellen Skalen

Abbildung 6 zeigt pro Skala des IIP den Prozentsatz statistisch signifikant verbesserter, verschlechterter und gleichgebliebener Personen. Im Gegensatz zur SCL-90-R (Kapitel 4.1.4.) wird hier jedoch von einer Irrtumswahrscheinlichkeit von 10% ausgegangen. Dies gründet sich auf den bisherigen Befunden, daß der IIP bzw. zwischenmenschliche Problembereiche eher änderungsresistent sind und deshalb die „Hürde für Veränderungen“ nicht so hoch gelegt werden soll wie bei den Symptomskalen. Insgesamt erkennt man, daß die meisten Personen selbst bei weniger konservativen Schranken für die Überschreitung der kritischen Differenz eher unverändert sind (zw. 55-89%). In Übereinstimmung zu den bisherigen Befunden ergeben sich für die Kontrollskalen PA, BC und NO deutlich mehr Verschlechterte als Verbesserte und für die Skalen „Introvertiert“, „Unterwürfigkeit“ und „Ausnutzbarkeit“ mehr Verbesserte als Verschlechterte.

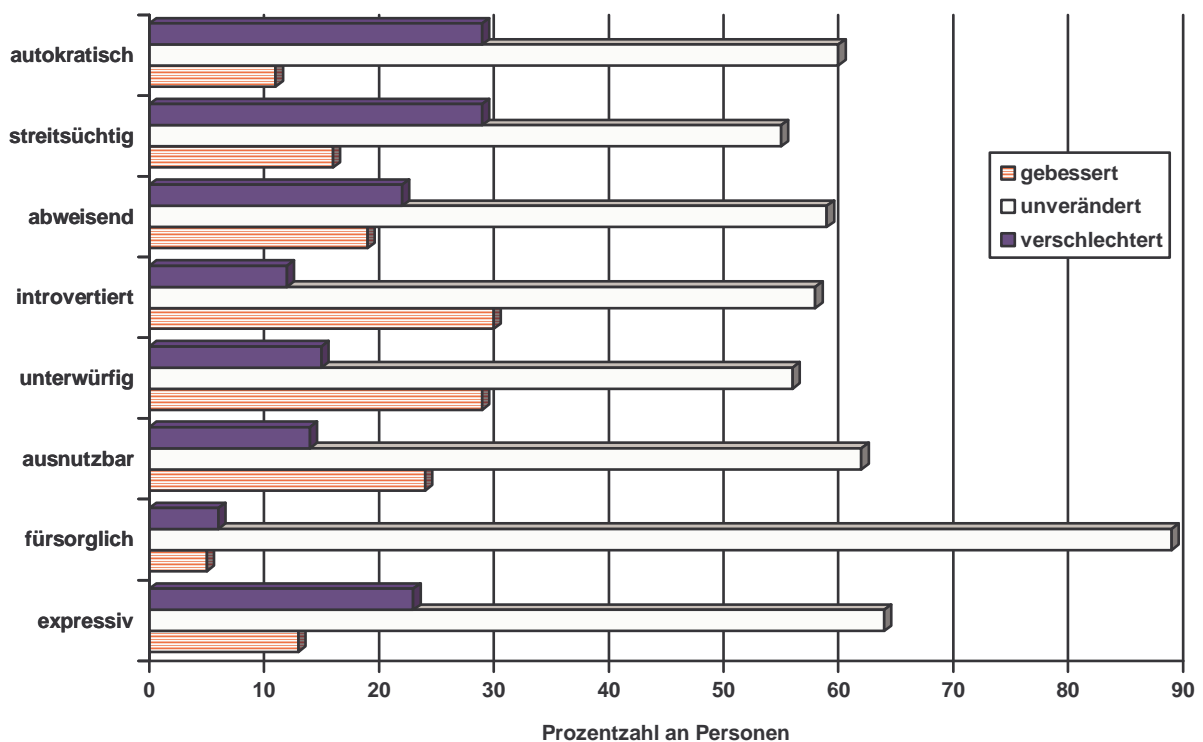


Abbildung 6: Prozentsatz statistisch signifikant veränderter Personen auf den interpersonellen Skalen.

5. Einschränkungen der vorliegenden Pilotstudie

- ◆ Die Veränderungen der Gruppenteilnehmer können nicht direkt auf die NIP Interventionen zurückgeführt werden, da die Möglichkeit besteht, daß sich vergleichbare Personen, die **nicht** mit NIP behandelt worden wären, im gleichen Ausmaß oder noch stärker verändert hätten. Es fehlen somit die Ergebnisse von unbehandelten Kontrollgruppen, um die erreichten Effekte wirklich auf die Intervention (NIP) zurückführen zu können.

Abhilfe: Zusätzliche Erhebung einer Kontrollgruppe, die hinsichtlich zentraler Merkmale (Alter, Geschlecht, Diagnosen, Einzugsgebiet...) mit der Therapiegruppe vergleichbar sein sollte. Diese Kontrollgruppe könnte aus Wartelisten rekrutiert werden, wobei kontrolliert werden muß, ob die Patienten in der Wartezeit noch zusätzliche Behandlungen erhalten haben (siehe z.B. Kazdin, 1994).

- ◆ Die Interpretation der Ergebnisse, daß die NIP Workshops die erwünschten Effekte (z.B. Verbesserungen auf den Symptomskalen) bewirkt haben, wird dadurch erschwert, daß sich die Teilnehmer häufig vor und nach den Workshops in ambulanter Therapie befanden.

Abhilfe: Es muß genau erhoben werden, welche Personen sich wie lange in zusätzlicher ambulanter Therapie befanden. Dann könnten alle NIP Teilnehmer mit versus ohne ambulanter Zusatztherapie miteinander hinsichtlich des Therapieerfolges verglichen werden. Ferner könnte man versuchen, den Therapieerfolg durch die Variable „ambulante Therapie“ (ja/nein; bzw. Dauer oder Intensität) vorherzusagen. Noch besser wäre es, vergleichbare ambulante Patienten mit und ohne zusätzlichen NIP Workshops hinsichtlich dem Therapieerfolg zu untersuchen. Damit kann die Frage beantwortet werden, ob ein zusätzliche NIP Kurs einen Zusatzgewinn für ambulante Therapie darstellt, ob dieser die ambulante Therapie verkürzt (Interesse der Kostenträger) oder intensiviert.

- ◆ Der Nachuntersuchungszeitraum von 1-2 Monaten ist sehr kurz. Kostenträger und auch die Patienten selbst haben ein Interesse an längerfristigen Psychotherapieeffekten. Hier haben sich gewisse Standards in der Katamnestik durchgesetzt. Z.B. empfiehlt die Deutsche Gesellschaft für Suchtforschung und Suchttherapie (1992, S. 121) als kürzesten Katamnesezeitraum ein halbes Jahr, gefolgt von den Intervallen 1, 2 und 5 Jahre. Ein kurzer Nachbefragungszeitraum hat den Vorteil, daß die Veränderungen zwar eher auf die Interventionen zurückgeführt werden können, aber den Nachteil, daß die Patienten „noch gar nicht richtig im Alltag stehen“.

Abhilfe: Halbjahres oder Jahreskatamnesen durchführen.

- ◆ Heterogenität der Teilnehmer. Dadurch, daß in den Workshops sowohl eher „gesündere“ Personen mit Selbsterfahrungsambitionen als auch wirklich behandlungsbedürftige „leidende“ Personen teilgenommen haben und auch hier untersucht wurden, und es versäumt wurde, diese Gruppenzugehörigkeit zu dokumen-

tieren, kann nichts über die wahre Effektivität für „echte“ Patienten ausgesagt werden. Durch das hier gewählte Vorgehen mit gemischten Personengruppen wird wahrscheinlich die „wahre Effektivität“ für Patienten *unterschätzt*, da die anderen Teilnehmer bei Beginn der Gruppen wahrscheinlich unauffälligere Werte haben und dadurch auch die Wahrscheinlichkeit einer Veränderung bis zum Nachbefragungszeitpunkt weniger wahrscheinlich ist.

Abhilfe: Die Diagnosen sind genau zu spezifizieren oder es ist zumindest der globale Schweregrad der Störungen bzw. Probleme genau zu dokumentieren (z.B. Globales Funktionsniveau aus dem DSM-IV oder der Beschwerde Beeinträchtigungsscore, BSS).

- ◆ Die Spezifität der Meßinstrumente ist fraglich. Standardisierte Skalen wie die hier eingesetzten Instrumente SCL-90-R oder IIP sind aus Gründen der Vergleichbarkeit mit anderen Studien von Vorteil. Ferner sind sie von den wissenschaftlichen Gütekriterien her gut abgesichert. Jedoch ergibt sich die Frage, ob man mit spezifischeren, auf die Ziele der NIP Methode genau zugeschnittenen Instrumenten nicht noch deutlichere Therapieerfolge hätte nachweisen können.

Abhilfe: Genaue NIP Ziele festlegen: In welchen Lebensbereichen erwartet man Verbesserungen beim Patienten? Im nächsten Schritt ist festzustellen, wie diese Ziele gemessen werden können. Diese Skalen (z.B. für emotionale Befindlichkeitsveränderungen; verändertes Selbstkonzept) bzw. auch selbstentwickelten Instrumente sollten dann zusätzlich zu den bereits erprobten Standardmaßen verwendet werden.

- ◆ Unklar ist auch weiterhin, welche einzelnen Elemente der Workshops gewirkt haben. Vielleicht sind dies völlig andere, als man erwartet hätte.

Abhilfe: Erfragen von Wirkfaktoren am Ende des Kurses oder bei der Nachuntersuchung (z.B. „Wie hilfreich war für Sie die Gruppenatmosphäre?“). Methodisch besser aber kaum durchführbar: Variation der einzelnen NIP Komponenten. Z.B. führt man eine Gruppe mit dem Element „Holding“ durch, die andere ohne „Holding“ (zufällige Zuteilung der Teilnehmer zu einer der beiden Bedingungen ist dann auch gefordert).

Literatur:

- Casriel, D. H. (1975). Die Wiederentdeckung des Gefühls-Schreitherapie und Gruppendynamik. München: Bertelsmann Verlag. (Original erschienen 1972: A scream away from hapiness)
- Deutsche Gesellschaft für Suchtforschung und Suchttherapie (Hrsg.) (1992). Dokumentationsstandards 2 für die Behandlung von Abhängigen. Freiburg i.Br.: Lambertus
- Frank, J.D. (1973). Persuasion and healing. Baltimore: Johns Hopkins Press.
- Franke, G. (1995). SCL-90-R - Die Symptom Checkliste von Derogatis - Deutsche Version, Testmanual. Weinheim: Beltz Verlag
- Grawe, K. (1992). Psychotherapieforschung zu Beginn der neunziger Jahre. Psychologische Rundschau, 43, 132-162.
- Grawe, K. (1995). Grundriß einer Allgemeinen Psychotherapie. Psychotherapeut, 40, 130-145.
- Grawe, K., Donati, R. & Bernauer, F. (1994). Psychotherapie im Wandel: Von der Konfession zur Profession. Göttingen: Hogrefe.
- Greenberg, L.S. & Safran, J.D. (1987). Emotions in Psychotherapy. NY: Guilford Press.
- Hartmann, A. & Herzog, T. (1995). Varianten der Effektstärkenberechnung in Meta-Analysen: Kommt es zu variablen Ergebnissen?. Zeitschrift für klinische Psychologie, 24 (4), 337-343.
- Hartmann, A., Herzog, T. & Drinkmann, A. (1992). Psychotherapy of Bulimia Nervosa: What is effective? A meta-analysis. Journal of Psychosomatic Research, 36, 159-167.
- Horowitz, L. M., Strauß, B. & Kordy, H. (1994). Inventar zur Erfassung interpersoneller Probleme (Deutsche Version, Testmanual). Weinheim: Beltz.
- Jacobson, N. S., Follette, W. C. & Revensdorf, D. (1984). Psychotherapy Outcome Research: Methods for Reporting Variability and Evaluating Clinical Significance.
- Kazdin, A. E. (1994). Methodology, Design, and Evaluation in Psychotherapy Research. In Bergin, A. E. & Garfield, S. L. (eds.): Handbook of Psychotherapy and Behavior Change (p 19-71). New York: John Wiley & Sons.
- Keller, W. & Schneider, W. (1993). Veränderungen interpersonaler Probleme im Verlauf ambulanter oder stationärer Gruppentherapie. Gruppenpsychotherapie, Gruppendynamik, 29, 308-323.
- Lueger, R. (1995). Ein Phasenmodell der Veränderung in der Psychotherapie. Psychotherapeut, 40, 267-278.
- Markert, P. & Mestel, R. (in Vorb.). Wirkung von New Identity Process Gruppen in stationärer Psychotherapie.
- McGaw, B. & Glass, G. V. (1980). Choice of metric for effect size in meta-analysis. American Educational Research Journal, 17, 325-337.
- Mestel, R. (1996). Stabilität des Therapieerfolgs von Patienten mit Persönlichkeitsstörungen verschiedenen Schweregrades in stationärer Psychotherapie. Vortrag auf dem 40. Kongreß der Deutschen Gesellschaft für Psychologie am 26. 9. 96 in München. Abstract im Tagungsband (Diskette).
- Mestel, R., Rudolf, A., Held, M. & Wittmann, W. W. (1996). Evaluation stationärer psychosomatischer Rehabilitation. Vortrag auf dem 6. Reha-wissenschaftlichen Kolloquium vom 3. bis 6. März in Bad Säckingen. Abstract im Tagungsband.

- Schmitz, B., Fydrich, T., Schifferer, E., Obermeier, L. & Teufel, J. (1996). Persönlichkeitsstörungen und Behandlungserfolg bei Patienten mit psychischen und psychosomatischen Störungen. In Schmitz, B., Fydrich, T. & Limbacher, K. (Hrsg.): Persönlichkeitsstörungen: Diagnostik und Therapie (S. 318-343). Weinheim: Psychologie Verlags Union.
- Tscheulin, D. & Schäfer, P. (1993). Differentielle Effektforschung an der Hochgrat Klinik: IV Effektstärkenberechnungen. Unveröffentlichter Bericht: Würzburg. Psychologisches Institut.
- Wittmann, W. W., Held, M., Rudolf, A. & Schulze, R. (1996). Gutachten über die Programmevaluations-Studie der Klinik für Psychosomatische Medizin in Grönenbach. Unveröffentlichter Bericht: Universität Mannheim.
- Wittmann, W. W. & Matt, (1986). Meta-Analyse als Integration von Forschungsergebnissen am Beispiel deutschsprachiger Arbeiten zur Effektivität von Psychotherapie. Psychologische Rundschau, 37, 20-40.
- Yalom, I.D. (1992). Theorie und Praxis der Gruppenpsychotherapie. Ein Lehrbuch. München: Pfeiffer Verlag (Original erschienen 1970: The theory and practice of group psychotherapy)