

8. Schaumburg C, Kächele H, Thomä H (1974) Methodische und statistische Probleme bei Einzelfallstudien in der psychoanalytischen Forschung. *Psyche - Z Psychoanal* 28: 353-374

CORNELIA SCHAUMBURG, HORST KÄCHELE, HELMUT THOMÄ, ULM

## **Methodische und statistische Probleme bei Einzelfallstudien in der psychoanalytischen Forschung\*** \*\*

*Übersicht:* Der Psychoanalytiker heilt, indem er forscht; dies „Junktim“ bedeutet aber nicht, dass psychoanalytische Falldarstellungen ( die als Prüfung psychoanalytischer Hypothesen interpretiert werden können) ohne weiteres wissenschaftlichen Kriterien genügen. Die vorliegende Arbeit orientiert darum über Möglichkeiten und Probleme der Anwendung empirisch-statistischer Verfahren auf psychoanalytische Einzelfallstudien.

### *Einleitung*

In der psychoanalytischen Verlaufsforschung bemüht man sich zunehmend darum, durch Tonbandaufzeichnungen von Behandlungen eine reproduzierbare Dokumentation des psychoanalytischen Dialogs herzustellen und somit empirischen Auswertungen eine verlässliche Basis zu geben (s. Kächele et al., 1973).

Zwar ist es „einer der Ruhmestitel der analytischen Arbeit“, dass Forschung und Behandlung bei ihr zusammenfallen (Freud, 1912, S.380), und man kann gewiss mit Freud (1926, S. 293/294) von einem „kostbaren Zusammentreffen“, einem „Junktim zwischen Heilen und Forschen“ sprechen. Aus diesem Junktim ergibt sich aber nicht ohne weiteres, dass klinische Falldarstellungen und Behandlungsberichte wissenschaftliche Kriterien erfüllen.

Es erscheint deshalb geboten, die Möglichkeiten und Bedingungen für die Anwendung empirischer und statistischer Methoden in Einzelfallstudien kritisch zu sichten. Da psychoanalytische Verlaufsforschung primär vom Einzelfall ausgeht, stießen wir bei unseren zunächst klinisch orientierten Auswertungen auf die Notwendigkeit, uns mit diesen empirisch-statistischen Problemen auseinander zusetzen. Die bei der Untersuchung eines einzelnen Falles auftretenden speziellen methodischen und statistischen Fragen sind insoweit kontrovers, als in der Statistik im allgemeinen große Fallzahlen gefordert werden, weshalb man lange Zeit Einzelfallstudien für unwissenschaftlich hielt. Wir sehen es deshalb als unsere Aufgabe an, besonders jene Seiten der Statistik darzustellen, die für Einzelfallstudien relevant sind.

### *Historischer Überblick*

Unser Literaturstudium führte zu der Feststellung, dass der Einzelfallstudie in den Anfängen der Psychologie und der Psychoanalyse größte Bedeutung zukam. Claude Bernard unterschied 1865 gruppenzentrierte von individuenzentrierter Forschung, welche für bestimmte, im einzelnen Individuum ablaufende Prozesse die einzig angebrachte Methode sei (zit. nach Shapiro 1964 ). Will man auf den Wert der Einzelfallstudie für die heutige Wissenschaft eingehen, so erscheint es sinnvoll, sich dies noch einmal ins Gedächtnis zu rufen.

---

\* Aus der Abteilung für Psychotherapie (Leiter: Prof. Dr. H. Thomä) der Universität Ulm.

Mit Unterstützung der Deutschen Forschungsgemeinschaft, Bad Godesberg.

\*\* Bei der Redaktion eingegangen am 30.5.1973.

In der Psychologie wurde eine Reihe fundamentaler Erkenntnisse an Einzelfallstudien gewonnen, was wir mit einigen Beispielen (nach Dukes, 1965) verdeutlichen wollen.

Im Jahre 1885 untersuchte Ebbinghaus an einer einzigen Versuchsperson, nämlich an sich selbst, Aspekte des Gedächtnisses beim Lernen von verbalem (sinnvollem und sinnlosem) Material und „schuf damit das Vorbild für den größten Teil der Untersuchungen über verbales Lernen in den letzten 80 Jahren“<sup>1</sup>.

1897 führte Stratton an sich selbst Wahrnehmungsexperimente mit Umkehrbrillen durch und zeigte, „dass die ‚absolute‘ Lokalisation retinaler Positionen wie oben–unten und rechts–links gelernt ist“ (Boring, 1942).

Ein im Lernvorgang häufig auftretendes Phänomen, das Plateau in der Lernkurve, entdeckten und erklärten Bryan und Harter 1899 anhand der Daten einer einzigen Versuchsperson.

„Den Bereich der Motivation betrifft ein Experiment der Physiologen Cannon und Washburn (1912), in dem Washburn selbst einen Ballon schluckte und bei dem dann gezeigt werden konnte, dass die kymographischen Messungen seiner Magenkontraktionen mit seinen introspektiven Berichten über Hungerempfindungen kovariierten.“

„In der affekt-psychologischen Literatur wurde Watsons und Rayners Studie (1920), in der darüber berichtet wird, wie die Furcht des kleinen Albert vor weißen Ratten konditioniert wurde, als ‚einer der einflussreichsten Artikel in der Geschichte der amerikanischen Psychologie‘ (Miller, 1960) begrüßt.“ Auf demselben Gebiet liegt die Bedeutung der von Jones 1924 durchgeführten De-Konditionierung der (nicht experimentell erworbenen) Furcht des kleinen Peter vor weißen, weichen Objekten. Dukes (1965) führt noch eine Reihe weiterer Beispiele an und schließt: „Eine weitere Darstellung der wichtigen Rolle der Einzelfallforschung in der Geschichte der Psychologie erscheint unnötig. Ein paar Untersuchungen, jede mit einer Wirkung wie ein Kieselstein, der eine Lawine ins Rollen bringt, sind der Anstoß für weiterreichende Entwicklungen in Forschung und Theorie gewesen. Andere haben eher wie fehlende Stücke eines fast fertigen Puzzle-Spiels gerade zur rechten Zeit Material für die Lösung der verschiedensten Streitfragen geliefert.“

Nicht nur in der traditionellen Psychologie, sondern gerade auch in der Psychoanalyse nimmt die Einzelfallstudie einen wichtigen Platz ein. Im Gegensatz zu Wallerstein und Sampson (1971) halten wir es nicht für „überflüssig, die außerordentliche Bedeutung der traditionellen (spezifisch psychoanalytischen) Fallstudien-Methode, die von Freud eingeführt wurde, noch einmal darzustellen. Das gesamte Schrifttum der Psychoanalyse ..., das die Phänomene normaler und abnormer Persönlichkeitsentwicklung und -funktion umfasst, bezeugt glänzend die Erklärungskraft der Theorie, die auf Daten aus dem Behandlungszimmer aufgebaut ist. Sie ist gewachsen in den Händen des Genius, der sie erfunden hatte, und derer, die nach ihm kamen, und hat eine wirklich außergewöhnliche Reihe von Einsichten in die Struktur der Psyche, in die Organisation psychischer Krankheiten, in die Kräfte, die in der Behandlungssituation wirksam sind, in die Prozesse psychischer Veränderungen und die Bedingungen der psychoanalytischen Technik geliefert.“ Dies lässt sich besonders an den bekannten Krankengeschichten Freuds illustrieren, die jede für sich einen Markstein im Verständnis eines speziellen Krankheitsbildes darstellen. Die zentralen Mechanismen des hysterischen Krankheitsbildes demonstrierte Freud zusammen mit Breuer in den „Studien zur Hysterie“ (1895) an mehreren Krankengeschichten (z. B. dem Fall der „Elisabeth von R.“); die aus den Analysen Erwachsener (z. B. des „Rattenmannes“, 1909) nur erschlossenen infantilen Neurosen konnte Freud in der Krankengeschichte des „kleinen Hans“ („Analyse der Phobie eines fünfjährigen Knaben“, 1909) direkt aufzeigen. Aber auch literarische Dokumente, wie z. B. die autobiographischen Aufzeichnungen des Senatspräsidenten D. P. Schreber (1903), verwandte Freud zur Aufhellung der Paranoia („Psychoanalytische Bemerkungen über einen autobiographisch beschriebenen Fall von Paranoia“, 1913).

---

<sup>1</sup> Alle Übersetzungen englischer Zitate stammen von den Verfassern.

Die großen Krankengeschichten Freuds hatten als Einzelfallstudien paradigmatische Bedeutung im Sinne von Thomas Kuhn<sup>2</sup>. Es wurden jeweils Entdeckungen über psychische Zusammenhänge gemacht, die gleichermaßen der Entwicklung der psychoanalytischen Methodik und dem Auf- und Umbau der psychoanalytischen Theorie dienten. Die Einzelfallstudien Freuds sind Dokumente einer „wissenschaftlichen Revolution“ (Kuhn, 1967). In dieser Phase der psychoanalytischen Forschung war es nicht vordringlich, die Reichweite der Theorien abzugrenzen, ihre Gültigkeit rigoros zu prüfen und sie mit alternativen Hypothesen zu vergleichen. Für die Psychoanalyse war eine Auseinandersetzung mit den methodischen Schwächen der rein klinischen Einzelfallstudie auch solange nicht zwingend, solange sie in der Pionierphase immer wieder , auf Neuland stieß. Auch in der Psychologie und Psychopathologie wurden diese Probleme nicht thematisiert, weil statistische Methoden, die im Einzelfall zur Anwendung kommen können und zur Validierung von Theorien geeignet sind, noch nicht entwickelt waren.

Der Einzelfallstudie kommt nunmehr eine große Bedeutung zu, weil die klinische und statistische Auswertung über die Theorieprüfung hinausführen und dort Neuland erschließen könnte, wo die klinische Einschätzung allein nicht ausreicht, vor Täuschungen zu bewahren. Systematische Einzelfallstudien, die eine formalisierte und systematische klinische Auswertung mit einer statistischen kombinieren, könnten geeignet sein, einen bestehenden Mangel in der psychoanalytischen Forschung zu beheben, weshalb wir uns nunmehr den speziellen Problemen systematischer Einzelfallstudien zuwenden wollen.

### *Pro und contra Einzelfallstudie*

Als Folge des oben erwähnten Mangels wurde zunächst in der Literatur allgemein die Ansicht vertreten, dass Einzelfallstudien für fundamentale Forschung nicht verwendbar seien (zit. nach Shapiro, 1961 b). Bellak und Chassan (1964) beschreiben das so: „Lange Zeit hat sich auf dem Gebiet der klinischen Forschung die Meinung gehalten, Vergleiche zwischen Gruppen von Patienten seien das sine qua non der statistisch validen wissenschaftlich-klinischen Forschung, und der Einzelfallstudie komme bestenfalls der Status der Intuition und der klinischen Einsicht zu, der für statistische Tests und Validierungsversuche nicht zugänglich sei... (Diese Meinung) hat unglücklicherweise diese von Natur aus oberflächliche Methodologie als einen einzigartigen Prototyp der Wissenschaftlichkeit in der klinischen Forschung... verewigt.“ Vorwiegend praktische Gründe führten jedoch zu einer Wiederentdeckung der Einzelfallmethodik, in deren Folge sich auch die theoretische und statistische Neubewertung abzeichnete. Besonders intensiv hat sich Chassan (1960) mit diesen Gründen auseinandergesetzt; im folgenden geben wir deshalb seine wichtigsten Argumente wieder. Ein Hauptproblem für den klinisch arbeitenden Forscher ist oft die geringe Anzahl von Fällen, die er in seine Untersuchungen einbeziehen kann. Dazu führen zu wenige für ein Projekt geeignete Patienten überhaupt, u.U. medizinische Kontraindikationen gegen die Einbeziehung geeigneter Patienten in eine Untersuchungsgruppe, Geldmangel und Mangel an Mitarbeitern. Bei der statistischen Auswertung von Ergebnissen kleiner unzuverlässig zusammengesetzter Gruppen ergeben sich dann folgende Schwierigkeiten: Zwar lässt sich von einem signifikanten Ergebnis (mit der Irrtumswahrscheinlichkeit  $\alpha$ ) sagen, dass es überzufällig ist; von einem nicht signifikanten Ergebnis jedoch lässt sich nicht sagen, dass es unbedingt zufällig sein muss.

Das Risiko erster Art,  $\alpha$ , das vom Forscher willkürlich festgelegt wird und die Signifikanz eines Ergebnisses direkt determiniert, gibt die Wahrscheinlichkeit dafür an, dass die Alternativ-Hypothese akzeptiert wird, obwohl die Null-Hypothese richtig ist. Dabei besagt die

---

<sup>2</sup> Als Paradigma bezeichnet Kuhn (1967), „dass einige anerkannte Beispiele für konkrete wissenschaftliche Praxis – Beispiele, die Gesetz, Theorie, Anwendung und Hilfsmittel einschließen – Modelle abgeben, aus denen bestimmte festgefügte Traditionen wissenschaftlicher Forschung erwachsen“.

Null-Hypothese, dass das Ergebnis zufälliger Natur ist; die Alternativ-Hypothese behauptet, dass das Ergebnis nicht zufällig, sondern bedeutungsvoll (= signifikant) ist. Daneben gibt es noch ein Risiko zweiter Art,  $\alpha$ , welches die Wahrscheinlichkeit dafür angibt, dass die Null-Hypothese akzeptiert wird, obwohl die Alternativ-Hypothese richtig ist. Dieses Risiko  $\alpha$  ist – bei gleichem Risiko  $\beta$  – in großen Stichproben kleiner als in kleinen. Bei der Arbeit mit kleinen inhomogenen Stichproben kann es also vorkommen, dass im Material vorhandene bedeutsame Unterschiede aufgrund des hohen Risikos zweiter Art nicht nachgewiesen werden können, d.h., dass das Ergebnis zwar sinnvoll interpretierbar wäre, aber statistisch nicht signifikant ist. Auch die Abhängigkeit der Beobachtungen voneinander, die in der klinischen Forschung fast die Regel ist, kann sich statistisch zuungunsten des Untersuchers auswirken. So ist z. B. die angenommene Unabhängigkeit der Beobachtungen von Patienten, die auf der gleichen Station liegen, durch die Kontakte der Patienten untereinander nicht gewährleistet.

Als Ausweg aus diesen Schwierigkeiten schlägt Chassan (1960) die intensive Untersuchung einzelner Patienten vor: „Es ist schwer zu verstehen, warum der intensive individuelle Ansatz in der formalen klinischen Forschung so sehr vernachlässigt worden ist. Man muss vermuten, dass der Grund für dieses Versäumnis ein generelles Missverständnis oder im günstigsten Fall eine Unklarheit über die Möglichkeiten ist, Schlussfolgerungen aus dem viel geschmähten Einzelfall zu ziehen. Es ist sicherlich noch nicht so allgemein bekannt, wie es sein sollte, dass die intensive statistische Untersuchung eines einzelnen Falles mehr bedeutsame und statistisch signifikante Information liefern kann als z. B. einmalige Beobachtungen eines Endzustandes an einer relativ großen Zahl von Patienten.

Dieser letzten Bemerkung stimmen sogar konservative Statistiker zu. So erklären Edwards und Cronbach (1952), dass der Informationsgewinn mit der Anzahl der unkorrelierten Variablen ansteige. Eine Verfeinerung der Messungen könne ebenso zu diesem Effekt führen wie eine Vergrößerung der Stichprobe.

Modelltheoretisch kann die Einzelfallstudie als „intensives Modell“ bezeichnet werden, dem die einmalige Messung an einer größeren Stichprobe als „extensives Modell“ gegenübergestellt wird (Chassan, 1960). Voraussetzung für die sinnvolle Untersuchung eines Einzelfalles ist, dass das untersuchte Merkmal innerhalb des Patienten über die Zeit hinweg variabel ist. Die Beobachtung der Variablen erfolgt unter verschiedenen Bedingungen. Die Randbedingungen, wie z.B. Alter, Geschlecht, Persönlichkeit und Vorerfahrungen des Patienten, bleiben relativ gleich und sind daher viel besser kontrolliert, als es bei einer größeren Stichprobe je möglich wäre. Andererseits sind die Randbedingungen dem Forscher durch seine intensive Beschäftigung mit dem Fall sehr gut bekannt, und er kann sie in seine Untersuchung einbeziehen, wenn ihm dies sinnvoll erscheint, wie z. B. bei Besuchen, die der Patient bekommt, oder Veränderungen auf seiner Station. Auch Brenman (1947) weist darauf hin, dass es der klinische Forscher im Gegensatz zum Experimentator fast immer mit mehreren kovariierenden Variablen zu tun hat.

Als besonders positiv stellt Chassan (1960) die Möglichkeit heraus, die Einzelfallstudie dynamisch gestalten zu können, das Design im Verlauf der Untersuchung verändern, Nebeneffekte berücksichtigen und weitere Fragestellungen einführen zu können, was bei einer größeren Stichprobe nur mit erheblichem Aufwand möglich ist. Die Einzelfallstudie ermöglicht eine Aussage über die untersuchte Variable bei dem untersuchten Patienten. Bei der Untersuchung einer größeren Stichprobe dagegen lässt sich – selbst bei Vorliegen eines signifikanten Ergebnisses – nichts über den Beitrag des einzelnen Patienten zu diesem Ergebnis, nichts über die Veränderung der untersuchten Variablen bei diesem Patienten sagen, was für den klinischen Forscher, der sich von seiner Untersuchung auch Hinweise für eine bessere Behandlung des Patienten erwartet, ein schwerwiegender Nachteil ist.

Auch Baldwin (1945/46) hebt hervor, dass z. B. Gruppenkorrelationen keinen oder nur ungenügenden Aufschluss über die intraindividuellen Korrelationen der Gruppenmitglieder geben.

Die ethische Seite betonen neben Shapiro (1957) Svenson und Chassan (1967) mit der Bemerkung, dass die Einzelfallstudie „glücklich die Lücke schließt zwischen klinischer Forschung für das Allgemeinwohl und klinischer Forschung für das Wohl des einzelnen Patienten, der in ihr involviert ist“. Wallerstein und Sampson (1971) kritisieren an dieser Bemerkung allerdings, dass sie „einseitig und deshalb irreführend nur die möglichen Vorteile, nicht aber die möglichen Schädigungen erwähnt“. Chassan (1957) weist darauf hin, dass nicht nur zwischen, sondern auch innerhalb von Individuen eine große Variabilität bestehe, die erst durch die Einzelfallstudie erfasst werden könne. Sie eröffnet somit auch völlig neue Aspekte der Reliabilität. Die intraindividuelle Variabilität von Messwerten wird jetzt nicht mehr als Messfehler, sondern als Erhöhung der Präzision aufgefasst. So baut z. B. jedes Beobachter-Patient-Paar ein eigenes interpersonales Feld auf; die Gesamtheit dieser Felder gibt die intraindividuelle Variabilität des Patienten an. Unterschiedliche Werte zweier Beobachter für denselben Patienten müssen nicht auf mangelnde Übereinstimmung der Beobachter zurückzuführen sein, sondern können ihre Gründe in einem unterschiedlichen Verhalten des Patienten zu den beiden Beobachtern haben.

In der Psychotherapieforschung hat man immer davon auszugehen, dass ein interpersonales Feld besteht, das von einer Arzt-Patient-Beziehung zur anderen auch bei gleicher theoretischer Ausrichtung der Psychotherapeuten wegen der „persönlichen Gleichung“ (Hartmann, 1964/65) verschieden ist. Anders ausgedrückt: die Therapeuten-Variable spielt eine entscheidende Rolle; dadurch kompliziert sich die Theorieüberprüfung erheblich. Was bei einem „homogenen“ Material qua gleicher Diagnose im Interaktionsfeld herauskommt, hängt ganz wesentlich vom Akteur ab, vom Therapeuten.

Auch das viel diskutierte „Konsensus-Problem“ der Psychotherapieforschung hat in dieser Variabilität der interpersonalen Felder eine wichtige Wurzel. Die Funktion der Gegenübertragung im diagnostischen Prozess kann dann als konstitutives Agens für das interpersonale Feld angesehen werden, womit die verschiedenen Beurteilungen des gleichen Materials durch erfahrene Psychoanalytiker ihre Erklärung finden (Seitz, 1966).

„Eine offensichtliche Schwäche der Einzelfallstudie ... ist, dass sie keine Aussage darüber liefern kann, ob die herausgefundenen Beziehungen auf alle anderen, viele andere, einige andere oder keine anderen menschlichen Wesen zutreffen“ (Janis 1958).

Mit diesem Zitat haben wir die wichtigste Einschränkung der single-case-study angeschnitten. In den modernen empirischen Wissenschaften besteht die Forderung, dass Generalisationen zwischen Fällen auf größeren Stichproben beruhen müssen. Zwar sind nach Dukes (1965) Aussagen über die Gesamtpopulation immer nur auf relativ kleine Stichproben gegründet, doch gehen in diese Informationen über die interindividuellen Unterschiede ein, was bei der Einzelfallstudie nicht der Fall ist. Deshalb ist die Generalisierbarkeit von Ergebnissen solcher Studien prinzipiell beschränkt.

Aber Chassan (1960) kann als engagierter Verteidiger der Fallstudien-Methode diese Einwände nicht akzeptieren. Der untersuchte Fall, so schreibt er, entstamme hinsichtlich seiner Variablen einer bestimmten Population; spezifiziere man die Variablen, könne man das Ergebnis auf die Population mit denselben Variablen generalisieren. In diesem Zusammenhang betonen Jacobs et al. (1966), dass bei dieser Spezifizierung die Bedeutung der einzelnen Merkmale in Betracht gezogen werden müsse; sie betrachten als wichtigstes Merkmal für die Generalisierung die Krankheit des Patienten. Da die Variablen beim Einzelfall sehr viel besser bekannt sind als bei jeder Stichprobe, folgt daraus nach Chassan (1960), dass die Generalisierbarkeit beim Einzelfall eher besser ist. Die Validität der Rückschlüsse von den Ergebnissen der untersuchten Gruppe hängt von der Zufälligkeit der Auswahl der untersuchten Gruppe ab, die bei klinischer Forschung im allgemeinen nicht gegeben ist. Außerdem werden oft alle den aufgestellten Kriterien genügenden Patienten in die Untersuchung mit einbezogen, so dass überhaupt keine Population verbleibt, auf die zurückgeschlossen werden kann.

Auch Shapiro und Nelson (1955) nehmen an, dass „die systematische objektive Untersuchung einzelner Abnormitäten in einzelnen Fällen zu Ergebnissen führe, die relevant für das generelle Funktionieren des betrachteten Individuums sind und auf viele andere Individuen übertragen werden können“.

Es scheint jedoch so, als ob besonders Chassans Meinung über die Generalisierbarkeit des Einzelfalles ein wenig zu optimistisch sei und den Einwand von Dukes (1965) nicht berücksichtige; die Identifizierung des Krankengutes einer Klinik zum Zeitpunkt einer Untersuchung mit der Gesamtpopulation, auf die Rückschlüsse z. B. bezüglich einer Behandlungsmethode möglich wären, scheint etwas engstirnig zu sein. Allerdings gesteht Chassan an anderer Stelle (1961; Bellak und Chassan, 1964) ein, dass ein einzelner Fall nicht die optimale Basis für klinische Forschung sei, sondern von mehreren intensiv untersuchten Einzelfällen sicher übertroffen werde, womit er den erwähnten interindividuellen Differenzen Rechnung zu tragen scheint.

Dukes (1965) beschäftigt sich weniger mit den generellen Vor- und Nachteilen der Einzelfallstudie, sondern diskutiert die Bedingungen, unter denen diese Methode sinnvoll angewandt werden kann.

1. Wenn Einzigartigkeit erfasst werden soll, ist der Umfang der zur Verfügung stehenden Stichprobe nur 1. Aber auch, wenn innerhalb einer Population keinerlei Variabilität besteht, erscheint die Heranziehung weiterer Fälle redundant.

2. „Im Gegensatz zu seiner begrenzten Nützlichkeit beim Aufstellen von Generalisationen von ‚positiven‘ Ergebnissen ist ein N von 1 bei einem ‚negativen‘ Ergebnis so nützlich wie ein N von 1000 beim Ablehnen eines behaupteten oder angenommenen allgemeinen Zusammenhangs.“

3. Die Einzelfallstudie ist angebracht, wo (aus welchen Gründen auch immer) keine Gelegenheit besteht, mehr als ein Individuum zu beobachten.

4. Auch für „problemzentrierte Forschung“ ist der Einzelfall ausreichend zum „Formulieren von Fragen, Definieren von Variablen und Abgrenzen von Problemen“.

Eine abschließende Bewertung der Einzelfallstudie geben Jacobs et al. (1966): „Die Argumente und Gegenargumente für und gegen das intensive Modell können also weder so gesehen werden, dass sie es als besten und letzten Ansatz in der klinischen Forschung herausstreichen, noch so, dass sie es als unwissenschaftlich und wertlos herabwürdigen. Es muss verstanden werden als nützlich, wissenschaftlich sauber und speziell für bestimmte Arten von Forschungsfragen geeignet.“

### *Allgemeine Methodologie*

Alle genannten Autoren sind sich, auch wenn sie es nicht unbedingt explizit ausdrücken, darin einig, dass das, was sie unter Einzelfallstudie verstehen, nicht mehr mit dem identisch ist, was Freud darunter verstand.

Das ist darauf zurückzuführen, dass der derzeitige Zustand der psychoanalytisch-klinischen Theorie weniger umwälzende Neuentdeckungen erwarten lässt; vielmehr steht die Ausarbeitung, Verfeinerung und Bestätigung von unzähligen klinischen Hypothesen im Mittelpunkt, was Rapaport (1966) die Aufarbeitung der „Faustregeln“ nannte, die bis auf den heutigen Tag den Großteil des klinisch-therapeutischen Instrumentariums ausmachen. So schreiben Wallerstein und Sampson (1971) in der Einleitung zu ihrer Arbeit: „Das Ziel und das Hauptthema dieses Artikels ist einfach die Rechtfertigung (und Ausarbeitung) der Notwendigkeit, die klinische Fallstudienmethode zu formalisieren; denn sie ist das wichtigste Forschungsinstrument und der wichtigste Zugang zur Erfassung des therapeutischen Prozesses in der Psychoanalyse.“

Der experimentelle Ansatz ist zwar die korrekteste Methode zur Überprüfung von Hypothesen (Campbell, 1967); er impliziert aber eine gewisse Manipulierbarkeit des Untersuchungsobjektes, die in der klinischen Situation nicht oder nur unzureichend möglich ist. Aber auch eine klinische Situation ist formalisierten Verfahren sicher zugänglich. So betont Campbell (1967), dass auch außerhalb von Experimenten richtige Schlussfolgerungen möglich seien, dann nämlich, wenn eine Beobachtung mit einer Erklärung übereinstimmt und andere Erklärungsmöglichkeiten nicht vorhanden sind oder ausgeschlossen werden können.

Die psychoanalytische Behandlungssituation wird oft als quasi-experimentelle Veranstaltung bezeichnet. Es erscheint uns adäquater, sie mit Shakow (1960) als semi-naturalistisch anzusehen. Die angemessenen Kontrollmethoden sind demzufolge nicht experimentelle, sondern statistische und Beobachtungsmethoden. Diese Auffassung räumt der Methodik der Einzelfallstudie einen wichtigen Platz in der Psychoanalyse ein. Damit meinen wir nicht, dass in der Einzelfallstudie ein Verzicht auf Exaktheit geleistet werden soll, sondern dass solche Verfahren zur Anwendung kommen sollten, die dem untersuchten System angemessen sind.

Shapiro (1961 b) gibt drei Bedingungen an, unter denen eine Formalisierung erfolgreich sein und die Einzelfallstudie zu einem Instrument auch grundlegender Forschung machen kann.

1. *Experimentelle Kontrolle.* Dieser Ausdruck, der inzwischen schon zu einer Phrase geworden sei, „impliziert, dass die Bedingungen, die Veränderungen im Auftreten eines Phänomens im individuellen Fall bewirken, gut genug bekannt sind, um eine Vorhersage – in messbarer Form – über das Auftreten dieses Phänomens in Situationen zu erlauben, in denen sich das Individuum nie vorher befunden hat. Die Beschreibung solcher Bedingungen läuft im Endeffekt auf die Beschreibung einiger der Gesetze hinaus, die das Phänomen in diesem Individuum bewirken“.

2. *Übertragbarkeit.* „Wenn man im Einzelfall ein Gesetz herausgefunden hat, bestimmt man anschließend den Grad, in dem dieses Gesetz bei der Beschäftigung mit anderen Fällen nachgewiesen werden kann.“ Man geht dabei von der Erwartung aus, dass die positiven Ergebnisse der Einzelfallstudie auch bei vielen anderen Fällen wieder auftreten. Das Fehlen einer Gesetzmäßigkeit im Einzelfall kann jedoch das generelle Fehlen dieser Gesetzmäßigkeit nicht annähernd so gut vorhersagen. Diese Ansicht Shapiros stellt keinen Widerspruch zu Dukes' Ausführungen über die verifizierende und falsifizierende Potenz der Einzelfallstudie dar.

3. *Entwicklung adäquater Messmethoden.* Solche Methoden müssen die relevanten Variablen so messen, dass Vergleiche innerhalb des Patienten über die Zeit hinweg und Vergleiche zwischen Patienten möglich sind.

In diesem Zusammenhang sind wir mit Chassan (1957) der Ansicht, dass man in der klinischen Psychologie jenes verdinglichende Denken überwinden muss, welches sich nur mit einem gegebenen Status (z. B. Diagnosen, der Einteilung in normal und abnorm etc.) beschäftigt, den prozesshaften Charakter des seelischen Geschehens dabei aber nicht berücksichtigt. „Statistische Ergebnisse werden ... oft mit einem Beiklang von Endgültigkeit und Sicherheit interpretiert, der dem Geist und dem Zweck der Statistik widerspricht“ (Chassan, 1967). Es müssen neue Methoden angewandt werden, um dieser dynamischen Betrachtungsweise gerecht zu werden.

Denselben Punkt betonen auch Wallerstein und Sampson (1971). Zusammen mit anderen Autoren sehen sie die Gefahr, dass „die Hervorhebung des Exakten (Quantifizierbaren) überflüssigerweise das Offensichtliche und Triviale bestätigt“. Sie fahren fort: „Der ernsthafte Erforscher des psychoanalytischen Prozesses darf unserer Meinung nach dieses Dilemma ... nicht einfach akzeptieren und damit die Notwendigkeit, zwischen Relevanz und Objektivität zu wählen. Er muss wichtige Probleme mit so exakten Methoden angehen, wie sie sein

Scharfsinn und seine Beharrlichkeit zu entwickeln vermögen, immer in dem Bewusstsein, dass eben nach neuen Methoden gesucht werden muss, wenn etablierte Forschungsmethoden die wichtigen Probleme in seinem Arbeitsgebiet nicht erfassen können.“

Die wissenschaftliche Erforschung des Einzelfalles verlangt vom Kliniker, dass er seine Hypothesen über Struktur und Psychodynamik präzisiert. Diese Präzisierung zeigt, welche psychoanalytischen Annahmen die Behandlungsstrategie bestimmen. So kann z.B. festgestellt werden, welche psychoanalytischen Theorien in einer Behandlung von mehreren hundert Stunden vom Analytiker überhaupt zur Anwendung gebracht werden, bzw. welche neuen Einfälle („clinical hunches“, GAP, 1959) das Verständnis einer Situation bereichert haben. Diese wichtige Seite der formalisierten Auswertung kann der heuristischen, hypothesenbildenden Funktion der Einzelfallstudie ein Moment von Sicherheit hinzufügen.

Gründe für die relative Unfruchtbarkeit bisheriger Ansätze in der Psychotherapieforschung liegen u. a. in der Schwierigkeit, die für die Technik der Behandlung relevanten Therapie-Variablen zu isolieren und zu quantifizieren (s. d. a. Luborsky, 1969).

Wallerstein und Sampson (1971) fragen sich im Anschluss an dieses Problem der Wirkung der Forschung auf die Praxis und an die von Dukes (1965) angegebenen „Indikationskriterien“ für sinnvolle Einzelfalluntersuchungen, ob es einen Punkt gebe, an dem verlangt werden muss, dass die am Einzelfall gewonnenen Hypothesen an größeren Gruppen überprüft werden. Die Antwort auf diese Frage hängt ihrer Meinung nach mit Bakans (1955) Unterteilung in Lehrsätze des generellen Typs und Lehrsätze des Aggregat-Typs zusammen.

„Lehrsätze des generellen Typs behaupten etwas, das wahrscheinlich für jedes einzelne Mitglied einer bestimmten Klasse wahr ist. Sie werden zunehmend gestützt durch jedes neue positive Ereignis, obwohl sie nie in einem formalen Sinn bewiesen werden können. Bei Auftreten des ersten negativen Ereignisses werden sie entweder in toto verworfen, oder – was wahrscheinlicher ist – die Klassengrenzen werden erneut definiert, was zu einer engeren Definition führt, die das negative Ereignis ausschließt und den neuen (eingeschränkteren und daher präziseren) Lehrsatz bestätigt. So nähert man sich der ‚Wahrheit‘ Schritt für Schritt an – durch die positiven Ergebnisse von Einzelfällen. Die Situation ist natürlich eine völlig andere bei den Lehrsätzen vom Aggregat-Typ, die Verhältnisse behaupten, die wahrscheinlich nur für eine Klasse wahr sind, wenn man sie als Masse betrachtet, und wo zunehmende Genauigkeit und Signifikanz mit zunehmender Größe und Repräsentativität der Stichprobe einhergehen“ (Wallerstein und Sampson, 1971).

Die Entscheidung, ob die am Einzelfall gewonnenen Hypothesen an größeren Gruppen überprüft werden müssen, hängt damit vom Anspruchsniveau der untersuchten Behauptung im Rahmen der psychoanalytischen Theorie ab.

Im Zusammenhang mit den methodischen Bedingungen der Einzelfallstudie erscheinen noch zwei Gesichtspunkte als wesentlich.

Dukes (1965) schreibt: „Obwohl Generalisationen über Verhalten gleichermaßen auf adäquaten Stichproben von Subjekten und Situationen beruhen, beziehen sich Gedanken über Stichprobenverfahren meist nur auf Subjekte. Entsprechend wird der Ausdruck ‚ $N = 1$ ‘ in der vorliegenden Diskussion gebraucht, um die ‚reductio ad absurdum‘ beim Ziehen von Stichproben von Subjekten zu bezeichnen. Er könnte sich jedoch ebenso gut (vielleicht noch besser, wenn man die Häufigkeit des Auftretens betrachtet) auf eine Einschränkung beim Auswählen von Situationen beziehen – z. B. auf ein einziges Labyrinth bei einem Lernversuch oder eine einzige Tapping-Aufgabe in einer Motivationsstudie. Diese beiden Möglichkeiten, Stichproben zu ziehen, im Sinn, spekuliert Brunswik (1956), der wichtigste Vorkämpfer repräsentativer experimenteller Designs: ‚In der Tat können saubere Stichproben von Situationen und Problemen letzten Endes wichtiger sein als saubere Stichproben von Subjekten, wenn man die Tatsache betrachtet, dass sich Individuen im ganzen wahrscheinlich ähnlicher sind als eine Situation der anderen.‘ Als Folge könnte der Ausdruck  $N = 1$  ebenfalls passend auf Stichproben von Versuchsleitern angewandt werden. Schon lange als eine mögliche Quelle der Varianz von Interviewergebnissen angesehen, ... wird der Untersucher



seit neuestem auch als eine Variable gesehen, die die Ergebnisse von Laboruntersuchungen beeinflussen kann.“

Wallerstein und Sampson (1971) bemerken, dass es keine Einzelfallstudie im engeren Sinne mehr ist, wenn ein Individuum als vollständige Population von Ereignissen aufgefasst wird, wobei die Ereignisse (nicht das Individuum) mit statistischen Verfahren bearbeitet werden können. In diesem Fall entfallen einige der Einschränkungen für statistische Schlussfolgerungen.

### *Spezielle Methodik*

In Einzelfalluntersuchungen werden – wie schon erwähnt – Daten erhoben, indem beim selben Patienten zu verschiedenen Zeiten dieselben Variablen gemessen werden. Die Daten liegen also in der Form von sog. Zeit-Serien vor. Selbstverständlich kann man Zeit-Serien nicht nur vom Einzelfall, sondern auch von einer Gruppe von Fällen erheben, was aber aus praktischen Gründen nur selten geschieht. Durch die bei der Untersuchung eines einzelnen Falles bestehende Möglichkeit, ihn zu verschiedenen Zeitpunkten unterschiedlichen Bedingungen auszusetzen, deren Einflüsse sich dann in der fortlaufenden Messung bestimmter Variablen niederschlagen, kann der untersuchte Patient zu seiner eigenen Kontrollperson werden. Viele Autoren betonen dieses Positivum, das den Forscher aller mit der Aufstellung einer Kontrollgruppe zusammenhängenden Schwierigkeiten enthebt.

Zeit-Serien-Analysen sind ein klassisches Experiment der Biologie und Physiologie des 19. Jahrhunderts und in diesem Bereich im Gegensatz zu dem der Sozialwissenschaften schon und noch immer anerkannt. Campbell (1967) steht auf dem Standpunkt, dass hier bisher oft die Frage nach dem Vorliegen von Veränderungen hinter der nach der Ursache dafür zurückgestellt worden sei, denn die Methodik zur Messung von Veränderungen sei noch wenig ausgebaut. Holtzman (1967) ist derselben Meinung, fügt aber hinzu, dass in den letzten Jahren, als Psychologen mit den damit zusammenhängenden Problemen bekannt wurden, statistischen Modellen für den Einzelfall immer mehr Aufmerksamkeit gewidmet worden sei.

Zur Erhebung und Auswertung von Zeit-Serien lassen sich prinzipiell die gängigen Verfahren benutzen, jedoch teilweise mit einigen Modifikationen und Einschränkungen. Im folgenden werden wir kurz auf einige dieser Methoden eingehen.

Eine in der klinischen Forschung allgemein und natürlich auch in der Einzelfallstudie häufig angewandte Erfassungsmethode ist die formalisierte klinische Beurteilung (Rating). Hierbei handelt es sich um eine sog. „Ad-hoc-Quantifizierung“ (Siegal, 1969) von Verhaltensweisen oder psychologischen Konstrukten, die einer direkten Messung nicht zugänglich sind. Neben den allgemeinen Reliabilitätsproblemen, auf die wir hier nicht eingehen wollen, ergibt sich bei klinischen Einstufungen über längere Zeiträume hinweg die Schwierigkeit, den der Beurteilung zugrunde gelegten (gedachten) Maßstab konstant zu halten. Deshalb ist zur Erhöhung der Reliabilität innerhalb und zwischen den Beurteilern eine sorgfältige Operationalisierung der zu quantifizierenden Verhaltenssequenzen und eine systematische Fixierung der Beobachtungsdaten erforderlich (Chassan, 1957).

Shapiro (1961 a) schlägt ein Verfahren vor, in dem nicht ein Beobachter, sondern der Patient selbst die Einstufungen vornimmt: seine Methode der Wahl ist der sog. persönliche Fragebogen. Er wird für jeden Patienten aus den Ergebnissen eines standardisierten Interviews einzeln konstruiert. Nach wiederholter Anwendung werden die Ergebnisse der einzelnen Messungen verglichen. Eine exakte Beschreibung dieser Methode gibt Chassan (1960).

Baldwin (1942) demonstriert eine ähnliche Methode, die er „personal structure analysis“ nennt, eine Methode, die – wie er betont – objektiv ist, aber auch Aspekte des Fallmaterials widerspiegelt, die sich der objektiven Erfassung entziehen, ähnlich wie z.B. die Methoden zur Auswertung des Rorschach-Tests. Das zugrunde liegende Material können Beschreibungen,

Briefe o.ä. sein. Beachtet werden die Häufigkeiten von Aussagen und das gemeinsame Vorkommen bestimmter Aussagen.

Als Annahmen gehen in diese Methode ein,

1. dass die Häufigkeit bestimmter Aussagen im Material ein Maß für deren Bedeutung in der Persönlichkeit ist, und

2. dass das überzufällig häufige gemeinsame Auftreten zweier Aussagen einen Zusammenhang in der Persönlichkeit anzeigt.

Die Aussagen werden in ein Kategoriensystem eingeordnet, das für jeden Patienten neu aufgestellt werden muss, damit die Individualität der Subjekte erfasst werden kann. Seine Aufstellung wird in Baldwins Arbeit eingehend beschrieben, ebenso die Auswertung mit Hilfe von Kontingenzkoeffizienten. Die personal structure analysis hat zwei Aufgaben:

1. das Aufdecken neuer, bis dahin unbekannter Zusammenhänge und

2. das Prüfen von Hypothesen.

Die personal structure analysis bietet einmal eine Ergänzung zum klinischen Interpretationsvorgang, zum anderen kann sie als eigenständige Methode einem Beobachter, der das Fallmaterial nicht kennt, ein Bild der untersuchten Persönlichkeit vermitteln.

Eine weitere Methode zur Erfassung von Veränderungen innerhalb eines Patienten ist die wiederholte Durchführung der Q-Sortierung (Stephenson, 1935; 1936). Dabei muss der Patient jedesmal eine Reihe von Feststellungen (z. B. „Ich bin oft ohne Grund niedergeschlagen.“) nach dem Grad ihres Zutreffens auf sich selbst ordnen. Die Ergebnisse der einzelnen Sortierungen lassen sich miteinander korrelieren; die Korrelationskoeffizienten sind ein Maß der Ähnlichkeit zwischen den Selbsteinstufungen des Patienten zu verschiedenen Zeitpunkten. Eine anschließende faktorenanalytische Verrechnung der Daten kann Aufschluss über die überdauernden Züge in der Selbsteinschätzung des Patienten geben.

Mit dem Semantischen Differential (Osgood, 1952), das von Hofstätter und Lübbert (1958) unter der Bezeichnung „Polaritätsprofil“ in den deutschsprachigen Raum eingeführt wurde, lässt sich der Bedeutungswandel von Begriffen im Erleben des Patienten erfassen. Die Aufgabe des Patienten besteht darin, bestimmte Begriffe (z.B. „Liebe“) assoziativ auf bipolaren Kontinua (z.B. „gut – schlecht“, „Mann – Frau“) einzustufen. Verbindet man die für einen Begriff angegebenen Skalenpunkte, erhält man ein Polaritätsprofil als Ausdruck der Bedeutung dieses Begriffes. Der Vergleich der Einstufung einer Reihe von Begriffen durch den Patienten zu verschiedenen Zeitpunkten erlaubt eine Aussage über die erlebnismäßige Veränderung der Bedeutung der Begriffe für den Patienten. Faktorenanalytisch können die überdauernden Strukturen in der Bedeutung eines Sets von Begriffen erfasst werden.

Verwandt sowohl mit der Q-Sortierung als auch mit dem Semantischen Differential ist die Grid-Technik, bei der Subjekte (z.B. „Mutter“) nach dem Grad ihrer Ähnlichkeit mit bestimmten Begriffen (z.B. „Liebe“, „schön“) eingestuft oder geordnet werden müssen. Korrelationsberechnungen zwischen den einzelnen Einstufungen geben die Gleichheit bzw. Veränderung des Gefühlsgehalts der Begriffe an. Darüber hinaus wurde eine der Faktorenanalyse ähnliche Auswertungsmethode entwickelt, die auch eine graphische Darstellung der Lage der Subjekte zu den Begriffen ermöglicht. Eine Beschreibung der Grid-Technik findet sich z. B. bei Fransella und Adams (1966).

Unter den verschiedenen Verfahren zur Auswertung von Zeit-Serien ist die Berechnung von Korrelationen das am häufigsten verwandte. Baldwin (1945/46) gibt es als Methode der Wahl an und bezeichnet es im Gegensatz zu den gängigen Korrelationen (= Gruppenkorrelationen = Beziehung zwischen zwei Variablen einer Population von Individuen) als intraindividuelle Korrelation, die er als „Beziehung zwischen zwei Variablen zu verschiedenen Zeitpunkten im Verhalten eines einzelnen Individuums“ definiert. Eine angefügte Faktorenanalyse kann die Integration von Verhaltensbeobachtungen in ein Konzept der individuellen Persönlichkeitsstruktur ergeben. Die Abhängigkeit der Messungen lässt sich bearbeiten, indem man in die Faktorenanalyse die Zeit als Variable eingehen und bei der Rotation den ersten Faktor mit der

Zeitvariablen zusammenfallen lässt. Alle weiteren Variablen lassen sich dann unabhängig von der Zeit interpretieren. Die bei der Untersuchung des Einzelfalls angewandte Art der Faktorenanalyse wird als P-Technik (= viele Messwerte von einem Patienten) bezeichnet (Cattell, 1943). Sie wurde von Luborsky (1953) zum erstenmal in der Psychotherapieforschung eingesetzt. Allerdings setzt die Verwendung eines faktorenanalytischen Persönlichkeitsmodells ein statisches Verständnis der Persönlichkeitsstrukturen voraus. Es wundert daher nicht, dass Autoren wie Chassan, die mit einem wahrscheinlichkeitstheoretischen Verständnis der Person und ihrer Dynamik arbeiten, prinzipielle Vorbehalte gegen die Verwendung der Faktorenanalyse als Forschungstechnik haben.

Bei aller Aufmerksamkeit, die man solchen grundlegenden Entscheidungen über das adäquate mathematisch-statistische Modell zuwenden muss, darf nicht vergessen werden, was Baldwin in allen seinen Ausführungen betont, nämlich die Bedeutung des klinischen Verständnisses für die Wahl der Fragestellungen und für die Interpretation der Ergebnisse.

Das Beurteilungssystem sollte zwei Erfordernissen genügen: einerseits muss es die Eigenart des individuellen Falles abbilden können, weil das die besondere Stärke der Einzelfallstudie ist; andererseits sollte es nicht – auf Kosten der Vergleichbarkeit – nur auf diesen einen Fall anwendbar sein.

Eine Fülle von Gedanken und Methoden zur Auswertung von Zeit-Serien gibt Holtzman (1967) an.

In der Zeit-Serien-Analyse ist die Zeit die wichtigste Variable, der alle anderen untergeordnet werden müssen. Die Grundgesamtheit ist die „Lebensdauer“ eines Systems, die Stichprobe ist eine relativ kurze Zeit-Serie innerhalb dieser Lebensspanne. Das Ziel ist eine valide Generalisation von der Stichprobe auf das Gesamtsystem. Bei der Auswertung von Zeit-Serien ergeben sich drei Schwierigkeiten:

1. Neben der Hauptvariablen Zeit gibt es noch andere Variablen, die berücksichtigt werden müssen.
2. Das Zeitintervall zwischen zwei Beobachtungen ist willkürlich und damit auch die Zahl der Beobachtungen.
3. Aufeinanderfolgende Messungen sind meist nicht unabhängig voneinander, so dass eine Reihe von statistischen Verfahren keine Anwendung finden kann. Die Abhängigkeit lässt sich jedoch durch die Vergrößerung der Intervalle zwischen den Messungen und durch geeignete statistische Verfahren reduzieren.

Trotzdem müssen die Messungen oft genug erfolgen, damit man überhaupt von einer Zeit-Serie sprechen kann. Zeit-Serien werden meist graphisch dargestellt mit  $n$  Beobachtungen auf der Abszisse und den Werten der Funktion auf der Ordinate. Sie zeigen typischerweise rapide Schwankungen zusammen mit einem langsamen, wellenförmigen Trend über einen längeren Zeitabschnitt.

Dieser Erscheinung liegen (wenigstens) drei Komponenten zugrunde:

1. ein Trend oder eine Bewegung über lange Zeit hinweg,
2. eine Schwingung in diesem Trend und
3. eine „zufällige“ oder Restkomponente.

Bei dem Versuch, diese Komponenten zu isolieren, gibt es noch immer ungelöste Schwierigkeiten. Besonders schwer ist die Unterscheidung zwischen dem Langzeittrend und den Schwingungen dieses Trends. Die oft angewandte Methode der gleitenden Mittelwerte kann diese Schwingungen artifiziell vergrößern.

Die Abhängigkeit sukzessiver Messungen kann mit Hilfe von serialen Korrelationen berechnet werden. Dabei wird jeder Wert zunächst mit dem folgenden, dann mit dem übernächsten, dann mit dem drittnächsten und so fort gepaart. Die Korrelationen pflegen mit der Entfernung der Werte voneinander (der „Verzögerung“) geringer zu werden. Die serialen Korrelationen können in einem Korrelogramm graphisch dargestellt werden. Auf der Abszisse wird dabei die Verzögerung und auf der Ordinate die Höhe der Korrelationen abgetragen. Meist ergibt

sich dabei eine abfallende Kurve. Wenn die serialen Korrelationen keinen Trend zeigen, handelt es sich um eine „stationäre“ Zeit-Serie; ihr Mittelwert und ihre Streuung sind von der Zeit unabhängig.

Betrachtet man die Veränderung nur einer Variablen mit der Zeit, so handelt es sich um eine einfache Zeit-Serie; bei der Betrachtung mehrerer Variablen um sog. multiple Zeit-Serien.

Die Berechnung von Korrelationen ist mit den gängigen Verfahren möglich, wobei gleichzeitig bestehende seriale Korrelationen bei der Signifikanzfeststellung berücksichtigt werden müssen.

Auch Campbell (1967) ist der Meinung, dass man bei der Überprüfung von Zeit-Serien auf signifikante Zusammenhänge oder Unterschiede Vorsicht walten lassen muss, da gerade die heterogene Abhängigkeit der Werte voneinander, bedingt durch die verschiedenen serialen Korrelationen, problematisch sei.

Statistische Lösungen für die Feststellung der Signifikanz des Unterschiedes zwischen zwei Messwerten eines Patienten und ähnliche Probleme geben Payne und Jones (1957).

Als die für Zeit-Serien-Analysen geeignetsten Variablen gibt Holtzman (1967) solche an, die sich beliebig oft messen lassen, geringe oder keine serialen Korrelationen aufweisen, aber genügend Variabilität, um interessant zu sein. Leider dürften bei den meisten Forschern andere als die hier angegebenen Gesichtspunkte eine Rolle bei der Auswahl der zu untersuchenden Variablen spielen. In diesem Zusammenhang warnt Holtzman (1967) vor der Benutzung der P- Technik der Faktorenanalyse, da dabei die serialen Korrelationen nicht berücksichtigt werden könnten.

Einen grundsätzlich anderen Ansatz als alle bisher genannten Autoren vertritt Chassan, der von wahrscheinlichkeitstheoretischen Überlegungen her eine Methodologie der klinischen Forschung und besonders der Einzelfallstudie entwickelt hat. Sowohl zwischen Individuen als auch innerhalb von Individuen besteht ein beträchtliches Maß an Variabilität, welches quasi die Schwingungen eines Systems um seine Ruhelage herum darstellt.

Von Mises hat bereits 1939 auf die Bedeutung der Wahrscheinlichkeitstheorie für die Psychoanalyse hingewiesen.

Jeder Schluss von einem Individuum auf ein anderes und von einer Verhaltensweise auf eine andere ist also mit einer gewissen Unsicherheit behaftet, hat aber auch eine gewisse Wahrscheinlichkeit. Die Wahrscheinlichkeit gibt an, wie stark ein Individuum oder ein Verhalten vom Durchschnitt abweicht. Mit zunehmender Zahl von Daten nimmt die Unsicherheit ab, aber Sicherheit wird nie erreicht. Auch in den nomothetischen Wissenschaften, z.B. in vielen Bereichen der Physik, sind nur Wahrscheinlichkeitsaussagen möglich.

Eine solche Wahrscheinlichkeitsbetrachtungsweise lässt sich – wie schon erwähnt – nicht mit einer Unterteilung in normal – abnorm und mit festen Diagnosen vereinbaren, sondern impliziert die Beschreibung des Patienten als multivariate statistische Verteilung und das Suchen nach stochastischen Modellen in der klinischen Forschung. Diese Tatsache sieht Chassan ebenfalls als einen Grund dafür an, die gängigen Reliabilitätsmessungen durch Korrelationskoeffizienten abzulehnen. Denn jede Reliabilitätsmessung setzt die Annahme einer grundlegenden Stabilität und die Begründung unterschiedlicher Messwerte durch Messfehler voraus. Da Chassan aber eine grundlegende Variabilität und die Determinierung jedes Messwertes durch Wahrscheinlichkeiten annimmt, bedeutet für ihn absolute Reliabilität absolute Vorhersagbarkeit, die nur bei absoluter Rigidität des Patienten möglich ist. Hohe Reliabilität von Tests kann vor diesem Hintergrund nur als Unempfindlichkeit interpretiert werden.

An diese von Chassan etwas zugespitzt formulierte Position lassen sich folgende Überlegungen anknüpfen. Hohe Reliabilität von Tests bezieht sich auf die konstanten Merkmale einer Persönlichkeit. Hieraus folgt: je höher die Reliabilität eines Tests, desto weniger ist er geeignet, die eher flexiblen Merkmale zu messen. Für Einzelfallstudien ist es deshalb entscheidend, Tests anzuwenden, die feine Veränderungen in Zeit-Serien abzubilden vermögen und zugleich für die qualitativen Seiten der Veränderungen sensibel sind.

### *Schlussbemerkungen*

Unser Ziel war es, im Rahmen einer orientierenden Arbeit deutlich werden zu lassen, welche empirisch-statistischen Möglichkeiten bestehen, die hypothesenprüfende Bedeutung klinischer Falldarstellungen zu verbessern.

Wenn sich die Psychoanalyse prinzipiell als Erfahrungswissenschaft versteht, muss sie auch den Schritt vollziehen, ihre Erfahrungen zu systematisieren und ihre Hypothesen am Material zu erproben. Hierfür sind gewisse Bedingungen in der Beobachtung, Registrierung und Auswertung notwendig, von denen wir einige zu diskutieren versuchten. Da sich psychoanalytische Verlaufsforschung immer im Junktum von Heilen und Forschen vollziehen wird, ist es unseres Erachtens entscheidend, wissenschaftliche Kriterien auf die Erfassung des psychoanalytischen Dialogs in seiner Veränderung in der Zeit anzuwenden.

### *Summary*

#### *Methodological and statistical problems of single case studies in psychoanalytic research.*

Psychoanalysts cure their patients while doing research work. This coincidence, however, does not imply that psychoanalytic case reports which may be understood as experiments validating psychoanalytic hypotheses automatically meet the requirements of scientific research. This paper discusses the potentialities and problems of empirical and statistical methods as applied to psychoanalytic studies of single cases.