

Qualitative und quantitative Methoden in der Psychotherapieforschung

Prof. Dr. Michael B. Buchholz

Vortrag am 21. April im Rahmen der
59. Lindauer Psychotherapiewochen 2009 (www.Lptw.de)

WAS TAUGT WAS?

Die Kontroversen zwischen qualitativen und quantitativen Forschern im Feld der Psychotherapie haben an Schärfe verloren, die Differenzen in der Sache werden nicht mehr so emotionalisiert geführt, auch weil man gemerkt hat, dass die einen nicht die guten Bewahrer und die anderen nicht die bösen Fliegenbeinzähler sind.

Verschiedene Einflüsse kamen hier zum Tragen. Da ist die Entwicklung von computergestützten Textauswertungsprogrammen wie Atlas.ti oder MaxQDA, mit deren Hilfe man einem Text vielerlei Codierungen zuweisen kann – deren Verteilungen dann einer quantitativen Berechnung zugänglich werden. Udo Kuckartz (2007), einer der Entwickler, sieht hier einen Grund für die Annäherung. Ein anderer Grund dürfte darin liegen, dass renommierte Forscher in Psychotherapie und Psychoanalyse verstärkt genaue Einzelfallforschung fordern und dabei verschiedene Textsorten produziert werden, die irgendwie bewältigt werden müssen. Wer einen Einzelfall untersucht, muss z.B. Berichte der einzelnen Stunden schreiben, aber auch genaue Audio-Transkripte anfertigen. Diese und weitere Textsorten müssen aufeinander bezogen werden und das kann qualitativ oder quantitativ geschehen, etwa durch Einbeziehung von geschulten Ratern. Schließlich ist es der Erfolg der quantitativen Forschung, durchaus auch zugunsten der Psychoanalyse, der ihr Recht gibt, denn nichts ist so erfolgreich wie der Erfolg.

Mir will es manchmal so vorkommen, als frage man, ob für das Erreichen eines Ziels Boot oder Fahrrad praktischer sei und ignoriere dabei die Beschaffenheit der Wege, die man einschlagen will. Bedenkt man, dass es Wege gibt, für die man Kletterschuhe braucht und andere, die man überfliegend abkürzen kann, dann merkt man, dass wir längst nicht mehr in der Welt der „Zwei Kulturen“ leben, die Charles Percy Snow 1959 ausgerufen hatte. Der Gegensatz von Geistes- und Naturwissenschaften hatte sowieso immer schon Technikwissenschaften, Jurisprudenz und große Teile der Medizin in ihr Schema nicht integrieren können. Wolf Lepenies hatte 1985 für die Soziologie einen eigenen Platz beansprucht und sein Buch „Die drei Kulturen“ genannt. Auf dieser Linie hatte John Brockman (1995) einen Band „The Third Culture: Beyond the Scientific

Revolution“ herausgegeben. Ich will an einigen Beispielen interne Schwierigkeiten qualitativer und quantitativer Forschung in der Psychotherapie so diskutieren, dass am Ende deutlicher werden könnte, worin tatsächlich ein Unterschied besteht und wie er genutzt werden könnte. Ich bin also nicht der Meinung, dass Unterschiede nivelliert oder überwunden werden sollten. Ich bin mit Gregory Bateson (1972) der Meinung, dass Unterschiede, wenn sie einen Unterschied machen, wertvolle Information darstellen oder erzeugen können.

ERSTES BEISPIEL

John Heritage und Douglas W. Maynard (2007) haben jahrelang medizinische Kommunikation untersucht. Wie sprechen Ärzte mit ihren Patienten? Wie werden schwierige Diagnosen mitgeteilt? Welche ethischen Konflikte werden im Gespräch wie gelöst? Diese Forscher fanden, dass die meisten Ärzte ihre Patienten nach 20 Sekunden unterbrechen und die Gesprächsinitiative übernehmen, indem sie Fragen stellen. Sie begründen das mit höherer Sachkompetenz und dass das Gespräch so verkürzt werden könne. So geführte ärztliche Erstgespräche dauern etwa 15 Minuten.

Heritage und Maynard haben nun Ärzte darin trainiert, erst dann selbst aktiv mit Reden zu beginnen, wenn der Gestaltbogen einer erzählenden Darstellung zu einem Abschluß gekommen ist und damit der Patient dem Arzt die Redeaktivität übergibt. Es zeigt sich, dass die meisten Patienten für die Darstellung ihrer Beschwerden dann ca. anderthalb Minuten benötigen, dann sind sie fertig. Wenn der Arzt jetzt Nachfragen stellt, verringert sich deren Zahl, das Gespräch selbst wird kürzer, aber die Zufriedenheit beider Beteiligten erhöht sich deutlich.

Am Ende fragen Ärzte immer, ob es noch weitere Probleme oder Beschwerden gebe? Das kann man im Englischen auf zweierlei Weise fragen: „Any more problems?“ oder „Some more problems?“ Aus dem Englisch-Unterricht werden Sie sich erinnern, dass die Wendung mit „any“ eher eine negative Tendenz impliziert, die Wendung mit „some“ eine positive Aufforderung hat. Ein Unterschied, der einen Unterschied macht. Die Autoren haben nun ermittelt, dass bei der Wendung mit „Some more problems?“ die Ärzte zweieinhalb mehr weitere Problemschilderungen erhalten. Der Arzt stellt an der Oberfläche eine Frage, macht aber mehr: er gibt Auskunft über *seine* Bereitschaft, weitere Probleme seines Patienten aufzunehmen oder nicht.

Ist das nun eine qualitative oder eine quantitative Studie? Ihrer Herkunft nach sind Heritage und Maynard qualitative Sozialwissenschaftler mit Spezialisierung in der sog. Konversationsanalyse, die sich die Erforschung nicht der Sprache, sondern des Sprechens zum Ziel gesetzt hat. Ein genereller Befund der konversationsanalytischen Forschungen ist, dass winzige Details im Sprechen große Unterschiede machen. Dafür ist die Studie über „Some“ und „Any“ ein recht eindrucksvolles Beispiel.

ZWEITES BEISPIEL

Konversationelle Winzigkeiten des Gesprächs können mächtige Wirkungen haben. Das gilt auch in umgekehrter Richtung, vom Patienten zum Therapeuten. Spence, Mayes und Dahl (1994) haben eine 600-stündige Analyse vollständig transkribiert und auf das sog. Anrede-Sie hin analysiert. Damit sind Wendungen von der Art gemeint, wenn der Patient seinen Analytiker direkt adressiert, etwa indem er sagt, „Ich denke gerade daran, wie Sie gestern gesagt haben...“ Es gibt nun Stunden, in denen ein solches Anrede-Sie sehr häufig ist, andere, in denen es gar nicht vorkommt. Die einen nennen die Autoren „related hours“, die anderen „isolated hours“. Das Anrede-Sie hat erhebliche Wirkungen. Im Durchschnitt machen Analytiker etwa 11 Äußerungen pro Sitzung. Aber in den related hours sprechen Therapeuten nicht nur mehr, sondern vor allem auch: früher! Nicht erst nach

15 oder 20 Minuten, sondern bereits direkt im Anschluss an das Anrede-Sie. Der Grund ist nicht schwer zu erraten: Wenn man direkt angesprochen wird, ist es schwer, sich zu entziehen. Aber ein solches Gesprächsmuster ist den meisten Analytikern wohl nicht bewusst, in den behandlungstechnischen Diskussionen darüber, wann man etwas sagt, kommt es jedenfalls nicht vor. V.a. zeigt diese Studie, wie ein Patient seinen Analytiker steuert und dessen Antwortverhalten beeinflusst.

DRITTES BEISPIEL

Wenn solche Details solche Wirkungen haben können, dann müsste das von guter quantitativer Forschung natürlich auch entdeckt werden. Davon handelt mein drittes Beispiel. Die Autoren der Studie von Raue et al. (1995) hatten zur Kenntnis genommen, dass das Arbeitsbündnis ein guter Prädiktor für therapeutischen Erfolg ist und nun wollten sie wissen, ob das Arbeitsbündnis von Therapeuten verschiedener Schulen auf die gleiche oder auf verschiedene Weise eingeschätzt wird. Verwendet wurde ein gut operationalisiertes Maß, die Working Alliance Scale (WAS), die der Einschätzung des Arbeitsbündnisses bei einer transkribierten Therapie dient. Therapeuten verschiedener schulischer Orientierung lag nun dasselbe Transkript vor, das sie nach den Vorgaben dieser Skala evaluieren sollten. Dasselbe Instrument also angewandt auf dasselbe Transkript, aber von Ratern verschiedener Orientierung. Einmal behavioral orientierte, dann psychodynamisch orientierte Rater. Es zeigt sich, dass die Beurteilungen des Arbeitsbündnisses nicht vom Transkript, sondern von der theoretischen Orientierung her am stärksten beeinflusst sind; dieser Faktor erklärt den größten Varianzanteil. Diese Untersuchung wird praktisch nicht zitiert. Sie wirft zu viele Fragen auf. Eine davon lautet, ob man überhaupt noch Rater für die Beurteilung von therapeutischen Prozessen verwenden kann, denn hier wird de facto selektive Wahrnehmung nachgewiesen, das Gegenteil von Objektivität. Schulische Vorliebe wirkt wie ein Vorurteil. Unsere Theorien machen gleichsam, dass wir sehen, was wir sehen und nicht sehen, was wir nicht sehen – basta. Die methodische Lösung, Rater zu schulen, greift nicht recht.

Denn man würde die Rater ja möglicherweise nur in der Vereinheitlichung ihrer Vorurteile schulen! Damit sie sehen, was sie sehen sollen – und weiterhin nicht sehen, was sie nicht sehen sollen. Die andere Antwort, gerade verschiedene Rater könnten ihre Vorurteile wechselseitig korrigieren, greift auch nicht recht. Mehrheitsübereinstimmungen haben ja noch nie Vorurteile beseitigt, sondern sie eher geschaffen. Immerhin aber muss man sagen, dass dieses Problem in aller Klarheit nicht etwa aus forschungsfeindlichen Motiven heraus erkannt wurde, sondern gleichsam im Kern der quantitativen Forschung selbst ermittelt wurde mit den besten verfügbaren Designs und Strategien. Ich meine, es ist ein Vorzug solcher Forschung, wenn sie präzise auf Probleme stößt, die sich aus ihrer eigenen Methodologie ergeben.

Das ist, ebenso wie das vorige Beispiel über „some“ und „any“ oder über das Anrede-Sie ein Beispiel aus der Prozessforschung. Offenbar kommt es auf feine Details an. Sie machen Unterschiede, die Unterschiede machen. Nicht die Unterscheidung von qualitativ und quantitativ ist das Problem, sondern dass nicht recht klar ist, was das Wort „empirisch“ bedeutet. Nicht umsonst stöhnen auch geschulte Rater sehr über den Zwang, etwas in eine kategoriale Schublade stecken zu sollen, was auch und gleichzeitig in andere Schubladen gehört. Sie leiden unter dem Gefühl, dass etwas wegfällt, was Teil der Empirie ist, aber irgendwie verschwindet.

Davon handelt mein viertes Beispiel.

VIERTES BEISPIEL

Die Umwandlung des natürlichen Gesprächs in wissenschaftlich prozessierbare Daten stellt also kein geringes Problem da. Es soll durch Methode, etwa den Einsatz geschulter Rater gelöst werden. Die dabei stattfindende Umwandlung von „Natürlichkeit“ in „Daten“ könnte wissenssoziologische Aufmerksamkeit beanspruchen, weil möglicherweise die beschriebenen Feinheiten in einer unbekanntem Weise unter den Tisch fallen.

Ich will nur *ein* Beispiel solcher Transformation aus einer Publikation (Schiepek, Genz, Schröder 1997, S. 240) zitieren. Die Autoren, der systemischen Familientherapie nahe stehend, wollten überprüfen, ob eine bestimmte Wirkung von zirkulären, hypothetischen und anderen Fragen tatsächlich zutrefte. Dazu haben sie als Datenbasis eine videographierte Paarbehandlung gewählt, also eine Einzelfallstudie. Unter der Überschrift „Methode“ beginnt dann die Beschreibung des Transformationsprozesses:

„Hauptdatengrundlage der Untersuchung waren die Videoaufzeichnungen der Paartherapie. Aus jeder Therapiesitzung wurden 3 Stichproben analysiert, die jeweils am Anfang, in der Mitte und am Ende der Therapiestunde lagen. Es werden jeweils Interaktionen im Anschluss an relativ klar auszumachende therapeutische Fragen untersucht. Der zeitliche Umfang der einzelnen Stichproben betrug 3-5 Minuten und richtete sich nach inhaltlich sinnvollen Einheiten, nämlich der jeweiligen Fragen und deren Beantwortung bzw. den Reaktionen auf die Frage. Die ausgewählten Stichproben wurden zum Teil transkribiert ... Danach wurden aus allen Stichproben Extensionen gebildet, entweder anhand der Transkripte oder direkt auf der Grundlage der Videoaufzeichnungen. Sowohl Transkripte als auch Extensionen stellten eine detaillierte Wiedergabe der verbalen (inhaltlichen) und nonverbalen (Körperhaltung, Mimik, Gestik, etc.) Aspekte der Kommunikation dar. In einem weiteren Schritt wurden die Extensionen kodiert. Dazu wurden die einzelnen Passagen in Verhaltens- oder Inhaltseinheiten sowie in Einheiten nonverbaler Kommunikation auseinanderdividiert.“

Beobachten läßt sich hier ein gestufter Transformationsprozeß. Die Paarbehandlung - das ist die professionelle Ausgangslage des „natürlichen“ Gesprächs. Dann gibt es davon eine Aufzeichnung - erste Transformation.

Deren Umwandlung in ein Transkript - zweite Transformation. Bildung von 3 Stichproben aus jeder Sitzung - dritte Transformation. Bildung von Extensionen aus den Stichproben - vierte Transformation. Verbale Beschreibung nonverbaler Interaktionen - fünfte Transformation. Usw. - Ähnliche Beschreibungen von Transformationsprozessen kann man nahezu allen wissenschaftlichen Publikationen im Feld entnehmen. Die Autoren dieser Studie erkennen freimütig an: „Eine solche Überprüfung von interaktionellen Vorgängen erfordert zwangsläufig Vereinfachungen eines komplexen Geschehens“ (S. 239). Auch solche Wendungen finden sich in vielen anderen Studien. Aber es ist nicht klar, ob man statt von Vereinfachung nicht konsequent von Daten-Vernichtung sprechen müsste.

Wenn man Literatur zu den verschiedenen Transkriptionssystemen liest, weiß man, welche Bedeutung Linguisten und Konversationsanalytiker diesem Problem widmen - was bedeutet das Heben der Stimme „an dieser Stelle“? Wie muss es notiert werden? Der Transkribent ist derjenige, der die eigentliche Datengrundlage für die wissenschaftliche Prozessierbarkeit schafft; wird nicht transkribiert, sondern wie in der psychotherapeutischen Forschung üblich, kategorisiert (etwa beim ZBKT oder dem SASB-Verfahren), dann sind es die Rater. Sie fungieren wie Transkribenten als „Augen der Wissenschaft“. Ohne sie sieht die Wissenschaft nicht.

Das Problem, das ich damit ansprechen möchte, ist das folgende: Jeder Psychotherapeut weiß, dass ein winziges Moment, etwa eine „wegwerfende“ Handbewegung am Ende einer Stunde oder das

verächtliche Zucken der Mundwinkel einem langen Gespräch eine völlig andere Bedeutung geben kann (Overlaet 1991). Wenn es solche kleinen Momente – „some“, „any“, das Anrede-Sie – sind, wie eigentlich kann man sie begründet weglassen, indem man ein langes Paargespräch auf 3mal 3 Minuten reduziert, die man dann untersucht?

Welche Bedeutung hat das Wort „empirisch“, wenn man soviel weglässt, von dem man nicht mal weiß, was es ist?

Fragen wir einmal anders: Was könnte die Grundlage einer solchen ja nicht seltenen Entscheidung sein, diesen oder jenen Ausschnitt aus dem Material für die weitere Analyse zu verwenden, andere hingegen nicht? Solche Entscheidungen werden auf der Grundlage einer *alltäglichen, einer natürlichen* Interaktionskompetenz getroffen. Empirische Prozeduren *stellen* auf diese Weise aktiv selektierend ihre Datenbasis *her* und das wirft Fragen bezüglich des Anspruchs an Objektivität auf. Sie können systematisch nicht wissen, was sie aus ihrem Untersuchungsansatz ausschließen. Sie können nur verfahren unter Einschluss dieses Ausschlusses, aber dieser Ausschluss bleibt gleichsam unbewusst oder unbemerkt. Wir sehen, auch quantitative Forschung beginnt qualitativ, muss so beginnen, das ist ihr Kern (Wolff 1994). Der Wissenschaftler kann erst handeln, wenn solche Transformationen vorgenommen wurden. Der quantitative transformiert in statistisch prozessierbare Einheiten, der qualitative Forscher transformiert in Transkripte oder prozessierbare Textsorten. Transformationen haben ihr Recht, sie heißen im wissenschaftlichen Kontext „Methode“ und die Wissenschaft kann nur solches Wissen anerkennen, das auf methodische Weise gewonnen worden ist. Gibt es eine empirische Lösung dieses Dilemmas? Ja, wenn man die Beobachtung auf eine Beobachtung höherer Ordnung erweitert.

Beobachtet man, wie ich es hier gerade tue, nicht nur Therapeuten bei der Arbeit, sondern Forscher, die ihrerseits Therapeuten beobachten, dann könnte sich etwas zeigen, was ein neues Licht auf das Verhältnis von qualitativen und quantitativen Forschungen wirft: Bei der Auswahl von relevanten Ausschnitten kommt eine natürliche Kompetenz hinzu. Forscher verstehen etwas von der Sache, sind mit ihrem Fall vertraut und wählen auf der Basis natürlicher Kompetenz aus. Ich will damit anregen, diese natürlichen Kompetenzen weit mehr zu untersuchen. Ich meine, der Wissenschaftstheoretiker Michael Polanyi (1958) hat recht, wenn er gegen Popper ausführte, dass die meisten Wissenschaftler nicht danach *streben*, ihre Theorie zu falsifizieren; sie nehmen es hin, wenn es so kommt. Jedenfalls die meisten. Aber im Grunde haben sie eine Intuition, ein Vorwissen von den Zusammenhängen und das gilt auch und gerade für Naturwissenschaftler. Meine Anregung ist also doppelt: a) Diese Intuition nicht länger aus dem Forschungsprozess zu verdrängen, sondern selbst wiederum sorgfältig und mit Feingefühl zu untersuchen; b) wissenssoziologisch zu untersuchen, nach welchen verborgenen Selektionskriterien Datenverlust oder Verdrängung zu dem beiträgt, was wir heute unter empirischer Forschung verstehen.

Qualitative Forscher haben das meist unter dem Titel der Selbstreflexion (Frommer 1989, 2000, Frommer und Rennie 2001, Poscheschnik 2005) gefordert. Merlin Donald (2008, S. 66f.) beschreibt, wie er als junger Kliniker beauftragt wurde, ein Instrument zu entwickeln, mit dem man genau festhalten konnte, wann Patienten nach einem Schlaganfall ihr Bewusstsein wieder erlangen. Es erwies sich als unmöglich, eine Liste mit Beobachtungskriterien aufzustellen oder experimentanalog vorzugehen, weil man solche Patienten auf keinen Fall beunruhigen darf. Es erwies sich als beste Strategie,

„auf den guten alten klinischen Sachverstand zurückzugreifen. Das heißt, die Fähigkeit erfahrener Kliniker, mentale Zustände zu erfassen, wurde zu unserem hauptsächlichen Messinstrument. Im Rückblick ist es durchaus bezeichnend, dass wir uns auf diese Strategie verlegen mussten. Die methodischen Notwendigkeiten hatten uns zu dem Eingeständnis gezwungen, dass das Urteil der Kliniker unseren experimentellen Paradigmen vier

entscheidende Punkte voraus hatte: Geschwindigkeit, Flexibilität, Feingefühl und Nuanciertheit.“

Und etwas später schreibt er:

“Das Erscheinungsbild des sich wieder einstellenden Bewusstseins ist keineswegs eindimensional, doch versierte Kliniker sind im Stande, sämtliche Indizien mühelos und simultan zu erfassen. Sie erkennen die Anzeichen dafür, dass der Patient zu Bewusstsein kommt, auf der Stelle, so schwach und indirekt sie auch sein mögen. Sie sind nicht die einzigen, die das können. Auch Angehörige des Patienten haben dafür ein feines Gespür.“
(Donald, S. 68)

Von Versuchen, solchen Subtilitäten auf die Spur zu kommen, handeln meine beiden letzten Beispiele.

FÜNFTES BEISPIEL

Wie durchaus subtile Dinge erforscht werden, zeigt eine Studie, die an 5 Frauen und einem Mann, alle als depressiv diagnostiziert, durchgeführt wurde. Ze'ev Frankel, Heidi M. Levitt, David M. Murray, Leslie S. Greenberg und Lynne Angus (2006) schreiben über „Assessing silent processes in psychotherapy“.

Schweigen kommt in „talk therapy“ weit häufiger vor als man im Allgemeinen berücksichtigt. Besonders die Autoren Greenberg und Angus sind in der qualitativen Narrationsanalyse ausgewiesen. Was macht Schweigen eigentlich aus? Wie unterscheiden sich verschiedene Arten des Schweigens und woran könnte man das erkennen? Schweigen ist nicht homogen, sondern aufteilbar in heilsames Schweigen und abwehrendes Verschweigen. Aber feiner noch geht's durchaus. Eine frühere Studie von Levitt ermittelte sieben Schweigetypen, die im PICS („Pausing Inventory Categorization System“) zusammen gefasst werden.

Damit werden Schweigepausen von mindestens 3 Sekunden Länge erfasst und anhand von kontextuellen, verbalen und paraverbalen Hinweisen folgendermaßen unterteilt:

- *disengaged silences*: hier zieht sich der Klient vermeidend von emotionalen Themen oder vom Prozeß zurück oder versucht, seine Gefühle durch das Schweigen zu kontrollieren - *Interactional silences*: wenn die Aufmerksamkeit des Klienten sich dem Therapeuten zuwendet, um das gerade besprochene Thema zu meiden - *Emotional silences*: wenn Klienten einem Gefühl nachspüren oder sich in ein Gefühl hineinzuversetzen versuchen - *Expressive silences*: wenn Klienten nach dem richtigen Wort für ein Gefühl suchen - *Reflective silences*: wenn erkennbar ein Stück Selbstanalyse vorgenommen wird, Verbindungen zu besprochenen Themen geknüpft oder Einsichten integriert werden.

Davon unterschieden wird noch das *mnemonic* und das *associational* Schweigen, wenn also jemand nach einer Erinnerung oder einem Einfall sucht. Die insgesamt 90 Sitzungen der 6 Behandlungen wurden komplett nach den verschiedenen Schweigeformen codiert und ergaben mehr als 1500 Schweigepausen. Es zeigt sich nun, dass produktives Schweigen in Therapie-Dyaden mit gutem Ergebnis häufiger vorkommt als die unproduktiven Varianten des Schweigens; bei schlechtem „outcome“ verhält es sich genau umgekehrt.

Das ist intuitiv einleuchtend, weil die Studie zwischen produktivem Schweigen und abwehrendem Verschweigen prägnant differenziert. Der Outcome wurde mit einigen Standardmaßen wie dem *Beck Depression Inventory* (BDI), der *Symptom Checklist-90* (SCL-90) und anderen gemessen. Reflexives Schweigen ist bester Prädiktor für positive Resultate, während obstruktives

Verschweigen mit negativen Behandlungsergebnissen einhergeht. Mich verblüfft, wie diese Befunde dann interpretiert werden:

“In this study, silences are understood as markers of internal processes, which are important because they suggest times in which clients’ internal experiences are so compelling that they disrupt the therapeutic interaction” (S. 621)

Wie kann das sein, so frage ich mich, dass man ein Phänomen, dessen Erfassung so offensichtlich mit einem *interaktiv* orientierten Instrument geschieht – nämlich an Kontexten vor und nach dem Schweigen – dann als „internal process“ zugerechnet wird? Meinen die Forscher hier, sie könnten diese inneren Prozesse sichtbar machen, während sie doch ein eindeutig interaktives Geschehen analysieren? Die Befunde sind hoch interessant und bestätigen die klinische Intuition, aber ihre Deutung ist unterkomplex. Denn im Behandlungszimmer sind es immer zwei, die schweigen. Damit Schweigen als „interner Prozess“ aufgefasst werden kann, damit ein Raum entsteht, in welchem das Schweigen produktiv für eine interne Klärung und Selbstklärung genutzt werden kann, muss eine *Zuschreibung* erfolgen; irgendwie müssen die Beteiligten sich darüber verständigen, wenn ein auftretendes Schweigen nicht unterbrochen werden sollte, weil da jemand die Stille für sein Nachdenken nutzen möchte. Erst aus einer solchen *gemeinsam* erarbeiteten Verständigung geht hervor, wem die beiden Beteiligten das Schweigen zurechnen wollen. Es wird dann nicht „*unser* Schweigen“, sondern „*dein* Schweigen“ – das kann, muss aber nicht therapeutisch fruchtbar sein. Dies mehr oder weniger stille, interaktiv aber *produzierte* Moment der *Zuschreibung* haben die Untersucher offensichtlich nicht in den Blick genommen. Es fällt zwischen die Maschen der Theorie und dann auch der empirischen Untersuchung. Wohin fällt es? Ist es etwa verloren? Die Antwort ist natürlich: nein! Es ist nicht verloren, es liegt in der verborgenen Kompetenz derjenigen, die das Schweigen codiert und geratet haben. Sie müssen diese Kompetenz als „native speakers“ mitbringen und nachvollziehen können, wem das Schweigen stillschweigend zugeschrieben wurde. Sie können dann einigermaßen gut (nämlich mit Interrater-Reliabilitätskoeffizienten von $r = 0.83$) codieren. Meine Überlegung zeigt, wie empirische Forschung sich subtil klinisch relevanten Momenten zuwendet, dass es aber ein überlagerndes Moment gibt, für das Kliniker ihre Sensibilität schulen. Es liegt in der Interaktion, in der therapeutischen Begegnung, wird aber in der Forschung eher als Rater-Kompetenz *genutzt*, nicht aber selbst untersucht. Schweigen im Behandlungszimmer ist immer kommunikative Angelegenheit. Eine Sache von zweien, erscheint aber in der Studie als Sache von nur einem, der Patientin oder dem Patienten.

SECHSTES BEISPIEL

Es gibt eine Studie, die auf subtile Weise etwas Ähnliches in den Blick nimmt, nämlich die Feindseligkeit zwischen Therapeuten und Patienten, deren Handhabung für Erfolg oder Misserfolg ziemlich ausschlaggebend ist, nicht aber für die Länge von Behandlungen; gerade verbissen Kämpfende können ziemlich viel Zeit miteinander verbringen.

Diese Studie, die sich mit gesprochener Konversation befasst, kommt aus Oslo von Anna Louise von der Lippe, Jon Trygve Monsen, Michael Helge Ronnestad und Dag Erik Eilertsen (2008). Hier interessieren Behandlungsfehler unter dem Titel: „Treatment failure in psychotherapy: The pull of hostility“. Kann man Erfolg oder Misserfolg einer Therapie aus dem Maß an feindseliger Aggression vorhersagen? Therapeuten sind da also beteiligt, was aber nicht heißt: schuld.

Forscher fanden “that hostile complementarity ... in treatment is rather common” (ebd.)

Das ist nicht erfreulich, aber bleiben wir gelassen. Die norwegischen Autoren gehen das schwierige Thema an, indem sie aus einer Studie von insgesamt 373 Patienten die kleine Anzahl von 28 auswählen. Genau die Hälfte hat einen positiven Behandlungsverlauf, die andere Hälfte einen negativen. Beide Gruppen von je 14 Patienten unterschieden sich diagnostisch nicht. Interessant ist, dass sie danach ausgewählt wurden, dass nur 14 Therapeuten beteiligt waren, die jeweils genau einen Patienten mit positivem *und* einen mit negativem Verlauf hatten.

Das entlastet, denn damit ist klar, dass es nicht „feindselige“ *Eigenschaften* des Therapeuten wären. Es geht vielmehr um die Qualität der therapeutischen Begegnung. Von den 28 Behandlungen, die mindestens 40 Sitzungen gedauert haben, wurden nun die ersten, die mittleren und die letzten 7 Minuten der dritten, zwölften und zwanzigsten Sitzung untersucht. Die Transkriptabschnitte wurden dann mit der SASB-Methode codiert, die ich hier nicht beschreibe (siehe Hartkamp 1994, 1997). Beobachtet wird, was geschieht, wenn z.B. ein Sprecher einen Angriff auf den Anderen richtet; dann ist es interessant, wie dieser reagiert – mit einer Gegenattacke oder mit Schweigen oder mit einer Äußerung, die als „affiliation“ codiert werden müsste. Und übt er damit Kontrolle aus oder behauptet er sich oder bleibt er emotional unabhängig und dennoch beteiligt? Wenn eine Äußerung mit einer Äußerung gleichen Typs beantwortet wird, dann heißt das entstehende Muster *Komplementarität* in der Sprache dieser Untersuchung und diese positiv oder negativ sein. Ein Beispiel für negative Komplementarität wäre, wenn einer dem anderen vorwurfsvoll sagt: „Du hast doch gerade gesagt, dass...“ und der andere reagiert mit „Du hast doch aber selbst gerade angefangen mit...“ Der eine Interaktionszug wird mit einem gleichen, negativ eskalierenden Interaktionszug beantwortet.

Die Ergebnisse sind klar. Feindselige Komplementarität war in der Gruppe der Misserfolge am größten, und zwar besonders in der 3. und 12. Sitzung, während sie in der 20. Sitzung abnahm. Freundliche Komplementarität konnte „positive Veränderung“ vorhersagen, ebenfalls in den Sitzungen drei und zwölf. Die Tendenz zu feindseliger Komplementarität war in der 20. Sitzung insgesamt weit weniger ausgeprägt. Ein anderes Muster zeigt, dass manche Patienten defensivabwehrend auf therapeutische Freundlichkeit reagierten und auch dann war das negative Ergebnis da! Die Autoren formulieren in einem schönen Bild

„that in successful therapy therapist and client follow each other, as in a dance (i.e., the overall balance of positive and negative affiliation can be predicted for one by knowing the other), while this harmony decreases over time in treatment failures” (S. 429)

Und etwas später schreiben sie zusammenfassend, es sei nicht möglich gewesen auf den zu zeigen, der die Saat des Misserfolgs im Dialog ausstreute. Es waren ja die gleichen Therapeuten in der Erfolgs- wie in der Misserfolgsgruppe. Die Befunde sind unabhängig davon, wer die negative Interaktion initiierte!

„ It seemed to be in the dialogue itself that the constructive or unconstructive therapeutic climate was created.” (Meine Kursivierung, MBB)

Die Folgerung ist klar: Therapeutische Arbeit besteht in mehr als Deuten, mehr als in “Intervention”. Sie muss maladaptive Muster herausfordern und sich ihnen aussetzen, sie muss schwierige Lebenseinstellungen korrigieren und die dabei aufkommende Verärgerung oder Wut aushalten, sie muss verzerrte Auffassungen deutlich machen und kann auch da streckenweise mit wenig Gegenliebe rechnen.

Aber es könnte sein, dass wir in Ausbildungen lehren und lernen könnten, genauer auf Anzeichen von Misstrauen, zurückgehaltenen Zweifel oder Kontaktabbrisse zu achten und v.a. auf die

Subtilitäten eigener gegenaggressiver Reaktionsneigungen, die natürlich oft „gut gemeint“ daher kommen. Der „pull of hostility“ ist groß und unsere Neigung, komplementär zu reagieren, nicht minder.

Wenn der Dialog, wie die norwegischen Autoren formulierten, unabhängig von Personen ist („It seemed to be in the dialogue itself...“) dann brauchen wir ein anderes Denken, dessen Gewinn darin liegen könnte, dass wir uns freier von Schuldzuweisungen diesen nicht seltenen klinischen Themen nähern. Wir sehen, dass ein Design, das nicht Personen, sondern Konversationen untersucht, solche Fragen auf eine ganz neuartige Weise zu beantworten vermag. Irgendwie liegt da etwas Wichtiges „in the dialogue itself“. Es verdient unser nachhaltiges Interesse.

SCHLUSSBEMERKUNG

Dieses, was da im Dialog selbst liegt, können wir bezeichnen. Es ist eine Art „natürliche Kompetenz“ analog der, wie Papoušek (1996) von der „intuitiven elterlichen Kompetenz“ spricht. Es ist meine grundlegende Auffassung, dass es so etwas gibt und dass therapeutische Ausbildung als Schulung und Verfeinerung einer solchen basalen intuitiven Kompetenz verstanden werden könnte. *Theorie* ist ein Wort, das weniger mit „wahr“ oder „falsch“ zu tun hat, sondern in der antiken Bedeutung verstanden werden könnte: als Instrument, etwas zu sehen.

So verstandene intuitive therapeutische Kompetenz wäre vor allem Wahrnehmungskunst, weniger „Interventions“- oder „Deutungs“-Kunst. Wer etwas sieht, der kann es auch ansprechen, wenn nur die grundsätzliche Bereitschaft, hilfreich sein zu wollen, relativ konfliktfrei gegeben ist. Aber bis man etwas sieht, müssen manche Brillen durchaus geputzt werden.

Warum spreche ich das an? Weil die hier referierten Befunde in meinem Verständnis alle in die Richtung weisen, dass es zutiefst menschlich ist, wenn wir aus Verhalten intuitiv auf Absichten schließen. Der kleine Säugling schaut auf den mütterlichen Zeigefinger, weil er die Geste noch nicht versteht; ist er erst älter als 9 Monate (Tomasello 2002), versteht er den Sinn der Geste und schaut dorthin, wohin der Finger weist. Verhalten dient ihm als Hinweis auf sinnhafte Absichten. Und später wendet er dieselbe Geste an und zeigt auf etwas, aber sein Blick schaut zur Mutter, ob sie auch dorthin blickt, wohin er ihren Blick dirigieren möchte. Die Psychoanalyse hat diese Idee, dass wir es immer mit sinnhaften menschlichen Intentionen zu tun haben, am Beispiel der Fehlleistungen um die Dimension der unbewussten Intention erweitert und Freud war bereit anzunehmen, dass wir in unserem alltäglichen Umgang alle „natürliche Psychoanalytiker“ seien. Also ohne dass wir seine Theorie gelernt hätten; wir tun's einfach. Wir verstehen, was ein Bundeskanzler insgeheim sagen wollte, als er öffentlich erklärte, am Ende käme es nur darauf an, „was hinten rauskommt“. Auch die nicht freudianisch informierte Republik lachte erheitert. Wir können nicht anders, wir müssen mit Sinn prozessieren. Nur Sinn schließt an Sinn an. Hier also kann der gesuchte Unterschied zwischen quantitativen und qualitativen Methoden in der Psychotherapieforschung gerade nicht liegen, auch wenn eine ganze Generation das gemeint hat. Auch quantitative Forscher müssen mit Sinnoperationen beginnen und unvermeidlich ihren Ergebnissen schließlich Sinn zuweisen. Der Unterschied liegt woanders.

Einmal, so meine ich, darin, ob wir Psychotherapieforschung oder Patientenbeforschung betreiben. Alle Forschung, die einen Patienten anhält, die Gespräche mit dem Therapeuten irgendwie zu benoten und die Veränderung des eigenen Zustandes zu protokollieren, ist Patientenbeforschung. Sie ist individualistisch in genau dem Sinn, als sie eine einzelne Person befragt. Sie ist mentalistisch insofern, als sie nicht Psychotherapie beobachtet, sondern Meinungen von Personen *über* Psychotherapie – und das ist um genau eine logische Ebene verschieden voneinander. Psychotherapieforschung hingegen kann nicht anders, sie muss das Geschehen *zwischen* mindestens

zwei Personen dokumentarisch festhalten, sei es per Video oder durch Audioaufzeichnung und analysieren, wie die Beteiligten aus ihrem sicht- und hörbaren Verhalten Sinn *produzieren* und Sinn einander *zuweisen*. Aus Sichtbarem und Hörbarem wird Erschließbares, das per se nicht sichtbar ist, aber aus der Abfolge der Redezüge erschlossen werden kann. Und qualitative und quantitative Beobachter können beobachten, wie Therapeuten ihrerseits Sinn aufnehmen und Sinn produzieren – und manchmal natürlich auch Un-Sinn.

Der zweite daraus folgende Unterschied ist der, mit welchen Dokumentationsmitteln die Psychotherapieforschung operiert. Die einen operieren mit Befragungen, Interviews von Therapeuten und Patienten, Verabreichungen von Fragebögen, diagnostischen Vorher- und Nachher-Wertungen usw. Dieser Typ von Forschung hat fraglos eigenen Wert. Wer aber *Psychotherapie* als aktualisierten (v.a. verbalen) Austausch untersuchen will, muss Transkriptionen nutzen. Und die wiederum können quantitativ oder qualitativ ausgewertet werden. Als Beispiel für quantitative Auswertungen von transkribierten therapeutischen Sitzungen möchte ich die Arbeiten von Erhard Mergenthaler (2008) nennen, als Beispiel für qualitative Auswertungen von transkribierten gruppentherapeutischen Sitzungen von Sexualstraftätern möchte ich auf unsere eigene Arbeit (Buchholz, Lamott, Mörtl 2008) verweisen. Lepper und Riding (2006) haben begonnen, von „transcript-based methods“ zu sprechen. Mir will scheinen, diese neuen Unterscheidungen zwischen Patienten- und Psychotherapieforschung und zwischen im weitesten Sinne Befragungen und Transkriptionen fasst die wissenschaftliche Art der Dokumentation viel besser und könnte die Diskussion befruchten. Die Forschung käme in die Lage, den Unterschied zwischen Forschung und Profession gewinnbringend zu nutzen (Buchholz 1999, 2006). Denn inwieweit Therapeuten Konzepte entwickeln, denen sie dann in ihrer Praxis folgen, bleibt ja immer erstmal eine offene Frage.

Und es könnte für die Forschung immer noch zu entdecken geben, *wie* sie das machen und, ob sie nicht vielleicht ganz anderen Konzepten folgen, vielleicht sogar ihren Intuitionen, deren Sinn nicht anders als aus ihrem Verhalten erschlossen werden müsste.

Literatur

Bateson, G. (1981): Ökologie des Geistes. Frankfurt/M.: Suhrkamp.

Buchholz, M. B. (1999): Psychotherapie als Profession. Giessen: Psychosozial.

Buchholz, M. B. (2006): Profession und empirische Forschung in der Psychoanalyse - ihre Souveränität und Integration. In: Psyche - Z Psychoanal, Jg. 60, S. 426–455.

Buchholz, M.B.; Lamott, F.; Mörtl, K. (2008): Tat-Sachen. Narrative von Sexualstraftätern. Giessen: Psychosozial

Buchholz, M. B.; Hartkamp, N. (Hg.) (1997): Supervision im Fokus - Polyzentrische Analysen einer Supervision. Opladen: Westdeutscher Verlag.

Buchholz, M. B.; Streeck, U. (Hg.) (1994): Heilen, Forschen, Interaktion. Psychotherapie und qualitative Sozialforschung. Opladen: Westdeutscher Verlag.

Donald, M. (2008): Triumph des Bewusstseins. Die Evolution des menschlichen Geistes. Stuttgart: Klett-Cotta.

- Faller, H.; Frommer, J. (Hg.) (1994): Qualitative Psychotherapieforschung. Heidelberg: Asanger.
- Frankel, Z.; Levitt, H. M.; Murray, D. M.; Greenberg, L. S.; Angus, L. (2006): Assessing silent processes in psychotherapy: an empirical derived categorization system and sampling strategy. In: Psychotherapy Research, Jg. 16, S. 615–626.
- Frommer, J. (1989): Möglichkeiten und Grenzen des systemtheoretischen Ansatzes in der Psychopathologie. In: Nervenarzt, Jg. 60, S. 65–70.
- Frommer, J. (2000): Qualitative Diagnostikforschung in der Psychotherapie. In: Psychother. Soz., Jg. 2, S. 203–223.
- Frommer, J.; Rennie, D. L. (Hg.) (2001): Qualitative Psychotherapy Research. Methods and Methodology. Lengerich: Pabst.
- Hartkamp, N. (1994): Sequentielle Strukturen in einer psychotherapeutischen Interaktion: Ergänzung oder Konkurrenz qualitativer und quantitativer Methoden. In: Faller, H.; Frommer, J. (Hg.): Qualitative Psychotherapieforschung. Heidelberg: Asanger .
- Hartkamp, N. (1997): Analyse einer Teamsupervision aus der Perspektive der SASB-Methode: Was läßt sich über latente Teamprozesse aussagen. In: Buchholz, M. B.; Hartkamp, N. (Hg.): Supervision im Fokus – Polyzentrische Analysen einer Supervision. Opladen: Westdeutscher Verlag .
- Heritage, J.; Maynard, D. W. (Hg.) (2007): Communication in Medical Care. Interaction between primary care physicians and patients. Cambridge: Cambridge University Press.
- Kuckartz, U. (2007): Einführung in die computergestützte Analyse qualitativer Daten. 2. aktualisierte und erweiterte Auflage. Wiesbaden: Verlag für Sozialwissenschaften.
- Lepenes, W. (1988): Die Drei Kulturen. Soziologie zwischen Literatur und Wissenschaft. Reinbek: Rowohlt.
- Lepper, G.; Riding, N. (2006): Researching the Psychotherapy Process. A Practical guide to Transcript-Based Methods. Hampshire/New York: Palgrave Macmillan.
- Lippe, A. L. von der; Monsen, J. T.; Ronnestad, M. H.; Eilertsein, E. (2008): Treatment failure in psychotherapy: The pull of hostility. In: Psychotherapy Research, Jg. 18, H. 4, S. 420–432.
- Mergenthaler, E. (2008): Resonating minds: A school-independent theoretical conception and its empirical application to psychotherapeutic processes. In: Psychotherapy Research, Jg. 18, S. 109–127.
- Overlaet, B. (1991): Interaction Analysis: Meaningless in the Face of Relevance. In: Int. J. Group Psychother., Jg. 41, S. 347–364.
- Papousek, M. (1996): Die intuitive elterliche Kompetenz in der vorsprachlichen Kommunikation als Ansatz zur Diagnostik von präverbalen Kommunikations- und Beziehungsstörungen. In: Kindheit und Entwicklung, Jg. 5, S. 140–146.

- Polanyi, M. (1958): Personal Knowledge. Towards a Post-Critical Philosophy. Chicago: University of Chicago Press.
- Poscheschnik, G. (Hg.) (2005): Empirische Forschung in der Psychoanalyse. Grundlagen - Anwendungen - Ergebnisse. Giessen: Psychosozial.
- Raue, P. J.; Putterman, J. T.; Goldfried, M. R.; Wolitzky, D. L. (1995): Effect of Rater Orientation on the Evaluation of Therapeutic Alliance. In: Psychotherapy Research, Jg. 5, S. 337–342.
- Schiepek, G.; Genz, W.; Schröder, I. (1997): Die differentielle Wirkung linearer, strategischer, zirkulärer und reflexiver Fragen in der systemischen Therapie. Einzelfallanalyse einer Paartherapie. In: Psychotherapeut, Jg. 42, S. 237–243.
- Snow, C. P. (1959): The two cultures and the scientific revolution. New York: Cambridge University Press.
- Spence, D. P.; Mayes, L. C.; Dahl, H. (1994): Monitoring the Analytic Surface. In: J. Amer. Psychoanal. Assoc., Jg.42, S. 43–64.
- St. Wolff (1994): Innovative Strategien qualitativer Sozialforschung im Bereich der Psychotherapie. In: Buchholz, M. B.; Streeck, U. (Hg.): Heilen, Forschen, Interaktion. Psychotherapie und qualitative Sozialforschung. Opladen: Westdeutscher Verlag .
- Tomasello, M. (2002): Die kulturelle Entwicklung des menschlichen Denkens. Zur Evolution der Kognition. Frankfurt: Suhrkamp.
-

Kontakt:

Prof.Dr.phil., Dipl.-Psych.
Michael B. Buchholz
Schlesierring 60, D-37085 Göttingen