

# Erzählen statt nur Zählen<sup>1</sup>

Peter Müller-Locher

Psychotherapie-Wissenschaft 8 (1) 47–52 2018  
www.psychotherapie-wissenschaft.info  
CC BY-NC-ND  
DOI: 10.30820/8242.09

*Zusammenfassung:* Der Autor zeigt, wie die Bemühungen nach hoher Reliabilität in diagnostischen Interviews Gefahr laufen, ihre Qualität des freien Erzählens zugunsten einer standardisierten Exploration der Patient/innen zu opfern und dadurch dem Zeitgeist des Zählens zu folgen. Als Beispiel dient ihm die Erfahrung als Schulungsverantwortlicher der Ratergemeinschaft in einer Forschungsstudie. Im Fokus steht dabei das Verfahren der Operationalisierten Psychodynamischen Diagnostik (OPD-2), in welchem eine hohe Einschätzungsübereinstimmung zwischen den unterschiedlichen Diagnostiker/innen zu erreichen versucht wurde.

*Schlüsselwörter:* Reliabilität, Beziehungsgestaltung, Mathematik, Literatur, Politik

Was hat Psychodiagnostik mit Politik zu tun? Was hat die Politik mit der Psychodiagnostik zu tun? Und welche Unterschiede sollen Erzählen und Zählen bezüglich psychologischer Diagnostik einerseits und hinsichtlich einer Politik psychologischer Diagnostik andererseits machen? Denn was haben das Erzählen von Geschichten als Literatur und das Rechnen mit Zahlen als Mathematik überhaupt mit Psychodiagnostik und was gar mit Politik zu tun?

Meine Ausführungen zu diesen prinzipiellen und weit ausgreifenden Fragen stützen sich auf konkrete und sehr begrenzte Erfahrungen, die ich als Psychodiagnostiker im Rahmen einer Psychotherapiepraxisstudie gemacht habe. Es handelt sich um die Praxisstudie ambulante Psychotherapie Schweiz (PAP-S). Sie ist ein inzwischen abgeschlossenes Forschungsprojekt der Schweizer Charta für Psychotherapie in Kooperation mit dem Klinikum der Universität Köln und dem Departement Angewandte Psychologie der Zürcher Hochschule für Angewandte Wissenschaften Zürich. 81 verschiedene Psychotherapeut/innen, die neun verschiedenen therapeutischen Schulen angehören, liessen während sieben Jahren die Therapieprozesse von 362 Patient/innen in einem gemeinsamen Forschungsdesign untersuchen (von Wyl et al., 2016).

Aus diesen Untersuchungen resultierte eine immense Fülle von Daten. Diese bestehen aus Fragebogen, ausgefüllt von den Patient/innen und den Psychotherapeut/innen, und aus Audioaufnahmen der meisten Therapiestunden. Alle Patient/innen wurden ausserdem zu Beginn und zum Abschluss der Therapie sowie ein Jahr nach Abschluss von Assessor/innen bzw. Diagnostiker/innen in sogenannte Assessment-Center eingeladen.

Während sechs Jahren durfte ich als Schulungsverant-

wortlicher dieser Rater/innen an der Studie mitwirken (vgl. Müller-Locher, 2016). In 40 halbtägigen Workshops versuchten mehr als 20 psychotherapeutische Rater/innen<sup>2</sup>, die als externe Diagnostiker/innen in den verschiedenen Assessment-Centern eingesetzt wurden, eine gute Übereinstimmung in ihren Befunderhebungen zu erreichen. Für die Diagnostik wurden sowohl diverse Fragebogen als auch die Strukturierten Klinischen Interviews SKID-I<sup>3</sup> und SKID-II<sup>4</sup> und die Operationalisierte Psychodynamische Diagnostik (OPD-2)<sup>5</sup> eingesetzt. Die Herausforderung, mit diesen Mess- bzw. Einschätzungs-

1 Der Titel meiner Reflexion nimmt einen Gedanken von Jonas Lüscher auf, der unserem Denken eine Tendenz des Zählens attestiert, wogegen er als Schriftsteller das Erzählen favorisiert. Denn das Erzählen negiere die Komplexität der Welt weniger als es jede mathematische Theorie tue.

2 Als Rater/innen wurden erfahrene Psychotherapeut/innen gesucht, die aus einer dritten Perspektive mit breit ausgewählten Diagnoseverfahren am Anfang und am Ende der Therapie und ein Jahr nach Therapieende die Therapieerfolge der an der Studie teilnehmenden Patient/innen untersuchen sollten.

3 Das SKID-I dient der Erfassung und Diagnostik ausgewählter psychischer Syndrome und Störungen ambulanter oder stationärer Patient/innen, wie sie im DSM-IV auf der Achse 1 («Klinisches Syndrom») definiert sind.

4 Beim SKID-II handelt es sich um ein zweistufiges Verfahren, bestehend aus einem Fragebogen und einem nachfolgenden Interview. Der Fragebogen dient dem Screening für die Merkmale der elf erfassten Persönlichkeitsstörungen. Im nachfolgenden Interview brauchen dann nur noch diejenigen Fragen gestellt zu werden, für die im Fragebogen eine «ja»-Antwort angekreuzt wurde.

5 «Die Operationalisierte Psychodynamische Diagnostik hat die Form eines multiaxialen Systems. Die OPD basiert auf den Achsen Krankheitserleben und Behandlungsvoraussetzungen (I), Beziehung (II), Konflikt (III), Struktur (IV) sowie psychische und psychosomatische Störungen nach dem Kapitel V (F) der ICD-10. Nach einem ein- bis zweistündigen Erstgespräch kann der Kliniker (oder der Forscher) die Psychodynamik des Patienten auf diesen Achsen einschätzen und in die Evaluationsbögen eintragen» (Arbeitskreis OPD, 2006 – Klappentext OPD-2). «Auch das neue System der OPD wurde während der letzten Jahre durch eine grosse Gruppe von psychodynamisch arbeitenden Psychotherapeuten aus den Gebieten der Psychoanalyse, Psychosomatik, Psychiatrie und psychologischen Psychotherapie entwickelt, die versucht, Operationalisierungen der therapielevanten psychodynamischen Aspekte zu formulieren» (ebd., S. 9). «60 Stunden Training (aufgeteilt in drei Kurse an drei Terminen) sind erforderlich, um das Manual reliabel anwenden zu können» (ebd., S. 8).

verfahren die teilnehmenden Patient/inn/en zu Beginn, am Ende sowie ein Jahr nach Abschluss ihrer Psychotherapie zu untersuchen, um die Wirkung ihrer Therapie auch aus externer Sicht einzuschätzen, war beträchtlich. Schliesslich sollten die Einschätzungen doch bei allen Erhebungszeitpunkten zu personenunabhängigen Befunden führen, und zwar in dreifacher Perspektive. Nicht nur sollten die befragten Patient/inn/en und ihre Therapeut/inn/en mit dem Resultat ihrer therapeutischen Bemühungen mehr oder weniger übereinstimmend zufrieden sein. Nein, auch der Anspruch, dass externe Einschätzungen aus dritter Sicht bei allen Zeitpunkten auch zwischen den diagnostizierenden Rater/innen übereinstimmen, sollte eingelöst werden können. Diese dritte Perspektive auf die untersuchten Therapieprozesse sollte gar eine besonders kritische Sicht auf die Wirksamkeit von den verschiedenen beteiligten Psychotherapieverfahren ermöglichen und deren Gültigkeit belegen. Als Voraussetzung für diesen Beleg an Validität sollte zwingend eine respektable Interraterreliabilität erzielt werden können. Das heisst, unabhängig von den jeweiligen Rater/innen sollten übereinstimmende Befunde bei jedem Erhebungszeitpunkt festgestellt werden können.

Die Ausgangslage war herausfordernd. Sie war aber auch prekär.

Die Rater/innen gehörten unterschiedlichen psychotherapeutischen Methoden und Verfahren an. Wie die an der Studie teilnehmenden Institutionen sprachen sie gewissermassen unterschiedliche Dialekte. Ihre praktische Erfahrung als Psychotherapeut/inn/en war ebenfalls sehr unterschiedlich und deckte ein weites Spektrum ab. Die Studie begann mit einem Pilotprojekt. Die erste Schulung der wenigen ersten Rater/innen in der anspruchsvollen Operationalisierten Psychodynamischen Diagnostik begann zeitgleich mit dem Start der Pilotphase. Zur zweiten Schulung nach zehn Monaten wiederum durch denselben zertifizierten Experten in der OPD stiessen 16 neue Rater/innen dazu. Nach den zwischenzeitlich selbst anberaumten und selbstorganisierten Übungsworkshops schritten die nun etablierten Rater/innen nach zwei Jahren zu einer ersten internen Messung der Interraterreliabilität. Nach weiteren drei Jahren und regelmässigen intervisorisch praktizierten Workshops organisierten sie gegen Ende der Studie eine zweite Messung der gefragten Einschätzungsübereinstimmung.

Zwischenzeitlich waren sechs Rater/innen ausgestiegen und vier neue ohne offizielle Schulung dazugestossen. Auch die Verteilung der durchzuführenden Assessments mit den Patient/inn/en und der dadurch entstehende Gewinn an Erfahrung blieben sehr unterschiedlich.

Die Prä- und Postinterviews wie auch die katamnestischen Assessments konnten ebenfalls nicht immer auf verschiedene Rater/innen verteilt werden. Interviewer-Verzerrungen konnten somit nicht ausgeschlossen werden. Die auf Video aufgenommenen Interviews von maximal 60 Minuten Dauer sollten den versammelten Re-Rater/innen verlässliche Aussagen und szenische Interaktionen liefern, um Struktur, Konflikt und Funktionsniveau des Patienten bzw. der Patientin übereinstimmend einschätzen zu können. Anders als in der normalen klinischen Praxis

war es im Studiensetting nicht möglich, in einem zweiten Interview fragliche Einschätzungen nochmals überprüfen zu können. Wurden den Rater/innen in den dem OPD-Interview folgenden Tests, zum Beispiel dem Strukturierten Klinischen Interview SKID, anhand geschilderter biografischer Episoden verständlicher, wie sie die Einschätzung in einer der OPD-Achsen präzisieren könnten, war diese Quelle der Erkenntnis den Re-Rater/innen, die sich nur auf das OPD-Video beziehen konnten, nicht zugänglich. Die Interviews sollten zudem gewissermassen auch didaktisch geführt werden. In der Video-Aufzeichnung sollte für alle teilnehmenden Re-Rater/innen gleichsam als beobachtende Zaungäste sichtbar und hörbar werden, woran die interviewführenden Diagnostizierenden gerade dachten. Das 60-minütige OPD-Interview sollte also nicht nur verlässliche Aussagen aus der biografischen und aktuellen Welt des Patienten bzw. der Patientin und szenische Interaktionen aus der Interviewbeziehung für eine Einschätzung liefern, sondern die weiteren Re-Rater/innen auch am Denken der Interviewführenden teilnehmen lassen.

Wie sollte aus so vielen Variablen bloss eine zuverlässige Einschätzungsübereinstimmung entstehen können? Sechs Jahre arbeitete die schliesslich verbliebene Gemeinschaft von einem guten Dutzend Psychotherapeut/inn/en auf dieses Ziel hin. Verglichen mit der publizierten Reliabilität der OPD blieben die Werte an Übereinstimmung eher bescheiden, in Zahlen festgehalten wenig bedeutsam.

Selbstkritisch gefragt: Was wäre alles erforderlich gewesen, um einen bedeutsameren zahlenmässigen Wert erreichen zu können?

Ressourcenorientiert gewendet: Was haben wir getan, um trotzdem einen mehrheitlich sehr bedeutsamen Prozess erlebt zu haben?

Zunächst jedoch zur Frage, wie denn ein psychodiagnostisches Testverfahren beschaffen sein muss, damit seine Reliabilität genügend, ja sogar hochstehend ist? Die einschlägige Literatur hält fest: Ein psychodiagnostisches Verfahren, so die allgemeinen Gütekriterien nach Lienert und Raatz (1998), muss objektiv, reliabel und valide sein. Die Ergebnisse müssen demnach unabhängig von Einflüssen der Untersuchenden oder der Untersuchungssituation bei der Durchführung, der Auswertung und der Interpretation zustande kommen. Solche Ansprüche sind am ehesten bei den psychometrischen Tests erfüllt. Schon weniger gut schneiden standardisierte psychologische Verfahren wie die strukturierten klinischen Interviews ab, welche zwar ausschliesslich deskriptive, symptomzentrierte Diagnosen erheben, aber bei der Befragung und Auswertung bereits einen grösseren Interpretationsspielraum öffnen.

Die Operationalisierte Psychodynamische Diagnostik wiederum setzt sich zum Ziel, die eigentlich unbefriedigende «symptomatologisch-deskriptiv orientierte Klassifikation psychischer Störungen um die grundlegenden psychodynamischen Dimensionen zu erweitern» (Arbeitskreis OPD, 2006, S. 27). Der OPD-Arbeitskreis entwickelte hierzu ein spezifisch konzipiertes OPD-Interview. In der offenen Eingangsphase wird dem Patienten in unstrukturierter Gesprächsführung Gelegenheit gegeben, sein

Verständnis seines Leidens in beziehungs-dynamischer Hinsicht mit dem Interviewer zu entfalten. Im folgenden mässig strukturierten Vorgehen wird nach biografisch relevanten Beziehungsepisoden zur Ermittlung von Konfliktmustern, dem Selbst- und Objekterleben und weiteren Struktureigenheiten wie der Selbst- und Beziehungssteuerung, der Kommunikation und der Bindung gefragt und die Patient/inn/en auch vertieft danach exploriert. Dieses im Prinzip mehrstündige diagnostische Verfahren, das «eine Synthese aus einem psychodynamischen Interview und einer Reihe von strukturierteren Interviewstrategien» (2006, S. 281) versucht, zeigt in den seither erforschten Gütekriterien gute bis sehr gute Übereinstimmungen in der Reliabilität, dies vor allem in der Strukturachse. Da die strukturellen Merkmale verhaltensnah definiert sind und wenig theoriegeleitet interpretiert werden müssen, zeigt speziell die Strukturachse eine gute klinische Praktikabilität und im Vergleich der Achsen die beste Interrater-Reliabilität.

Und wenn wir uns noch weiter wegbewegen vom verhaltensnahen Messen und Zählen, stossen wir noch tiefer in die Welt des bedeutungsvollen Erzählens vor. Die sogenannten narrativen Interviews stützen sich auf die Erzähltheorie und methodisch auf die biografisch-narrative Gesprächsführung. Will die Erzähltheorie noch Gesetzmässigkeiten von universaler Gültigkeit in einer systematischen Beschreibung der Darstellungsform eines Textes anstreben, so zielt das narrative Interview nicht mehr auf die Ermittlung objektiver Daten wie zum Beispiel eines Lebenslaufs. Vielmehr will das narrative Interview durch die Erfassung und Erzählung der Biografie des Interviewten dessen eigene Perspektive in Form der von ihm konstruierten subjektiven Sinnzusammenhänge zeigen. Zentral ist dabei die Annahme, dass die soziale Wirklichkeit jeweils erst im Rahmen kommunikativer Interaktionen hergestellt wird (vgl. hierzu Küsters, 2009 S. 18). Im freien Erzählen über selbsterlebte Ereignisse ergeben sich subjektive Bedeutungsstrukturen, die sich einem systematischen Befragen im Interview eher verschliessen würden. Narrative Kommunikation, so die These, sei eine spezifische Form der Selbstreflexion.

Was zeigt dieses breite Spektrum in Sachen Gütekriterien bei psychologischen Diagnoseverfahren auf? Von offenbar hochgradig erzielten Werten insbesondere der Reliabilität bei psychometrischen Tests bis hin zur individuellen Gewinnung subjektiver Bedeutungsstrukturen bei narrativen Interviews, bei deren Anwendung schon gar nicht mehr nach Reliabilität gefragt wird?

Wissenschaftstheoretisch eröffnet sich ein Spannungsfeld zwischen wissenschaftlichem Realismus einerseits, der von Objekten und einer Annäherung an deren Wahrheit ausgeht, und andererseits dem sozialen und radikalen Konstruktivismus, wonach auch scheinbar objektive naturwissenschaftliche Tatsachen das Ergebnis von Prozessen der sozialen Konstruktion seien und somit jede Wahrnehmung vollständig subjektiv werde.

Psychologische Diagnoseverfahren können Erfragtes und Beobachtetes zu Daten reduzieren und diese als erzielte Werte zählen und verrechnen. Narrative, das heisst

erzählende Interviews, die sich gerade nicht mehr als Diagnoseverfahren verstehen, öffnen Zugänge zum subjektiven Verstehen eigener Deutungen des Interviewten. Die Psychologie des Verstehens in der ganzen Vielgestaltigkeit der Hermeneutik unter Einschluss der Fremdheit des zu Verstehenden und der Intentionalität des Interviewers zeigt, dass das Kriterium der Überprüfbarkeit entfällt.

Je offener also ein Interviewter zum Erzählen eingeladen wird, desto geringer fällt jede Reliabilität aus. Je strukturierter und standardisierter ein Interviewter exploriert wird, desto höher wird das Gütekriterium der Reliabilität ausfallen.

Die Operationalisierte Psychodynamische Diagnostik versucht zwischen sich offen entfaltender Erzählung und strukturierender Exploration eine Synthese zustande zu bringen. Denn je strenger die Interviewer geschult werden, die strukturierenden Teile im OPD-Interview internalisiert anwenden zu können, und je strenger vor jeder offiziellen Interraterreliabilitätsmessung eine angeblich vorgängige selektionierende Befundunsicherheitseinschätzung vorgenommen wird (vgl. Arbeitskreis OPD, 2006, S. 60), umso höher würden die Zählwerte ausfallen.

Was wäre also angezeigt gewesen? Zuerst seriöse Ausbildung durch Experten mit nachfolgender Zertifizierung der Rater/innen und dann erst der Start der Studie?

Das wollte und konnte die Leitung der PAP-S Studie nicht leisten. Es war auch nicht nötig. Die Wirksamkeit der untersuchten Psychotherapien konnte auch durch andere Ergebnisse hinreichend bestätigt werden.

Wie aber ging die Rater/innengemeinschaft um mit der Vielzahl von Variablen, mit den realen Beschränkungen im Studiensetting und mit den Variablen, die im Vorhinein den geforderten Gütekriterien eines psychologischen Diagnoseverfahrens widersprachen?

In den von den Rater/innen selbst anberaumten OPD-Übungsworkshops, die sie während gut fünf Jahren durchführten, entwickelten sie bald eine hilfreiche Struktur.

«Ohne leitenden Experten, sondern in einer intervisio-rischen Zusammenarbeit bemühten wir uns grundsätzlich, voneinander zu lernen. Zwei Vorgehensweisen wurden praktiziert. Entweder sahen wir uns ohne Unterbrechung ein Interview an und machten uns ohne Diskussion gleich an unsere individuellen Einschätzungen. Die danach festgehaltenen Werte zeigten dadurch ungeschminkt das weite Spektrum unserer Einschätzungen auf. Das Resultat war oft ernüchternd, was wiederum längere klärende Diskussionen provozierte. Oder wir unterbrachen ein wiedergegebenes Interview und entwickelten in Kleingruppendiskussionen Hypothesen zur Einschätzung. Wir liessen uns dabei von zwei Fragen leiten: Was hat die Interviewführung ansatzweise schon zutage gefördert? Was müssen wir durch welche Exploration noch wissen, um am Ende zu einer verlässlichen Einschätzung zu gelangen? Mit dieser Vorgehensweise schärften wir einerseits unsere Wahrnehmung und förderten andererseits unsere Explorationsplanung. Aber nur zu oft stiessen wir dabei auf unsere verbreitete suboptimale OPD-Interviewführung» (ebd., S. 54).

In der Planung der Workshops einigten sich die Ratenden auf zwei Übungstermine pro Quartal à fünf Stunden. Die

Teilnahme an den Übungsworkshops war freiwillig und wurde nicht bezahlt. In ihrer demokratischen Selbstorganisation einigten sie sich auf eine Teilnahmeverpflichtung von 50 Prozent an den jeweils acht Workshops pro Jahr. Jede/r Rater/in musste in diesen Jahren zwei Interviews vorstellen: ein erstes, welches sie unter dem Aspekt der Interviewführung konsensuell rateten, und ein zweites, welches sie einer inoffiziellen Interraterreliabilitätsmessung unterzogen.

«Beim ersten Interview halfen uns bewusst würdigende und teils auch übertragungsfokussierte Fragen bei unseren Unterbrüchen im Abspielen der Videos. Nach der entwertend angekommenen Kritik des Experten zur Interviewführung wollten wir uns in unserer grossen Unterschiedlichkeit nicht auseinanderdividieren lassen. So interessierten uns Fragen, durch welche Besonderheiten interviewte Patienten die Interviewerin zu welchen Haltungen und Interventionen veranlasst haben könnte? Motiviert beispielsweise eine strukturelle Schwäche des Patienten die Interviewerin zu einer Vielzahl von strukturierenden Fragen? Oder beeinflusst eine tiefe Angst der Patientin den Interviewer zu einer schonenden Gesprächsführung?» (ebd., S. 56).

In diesen Diskussionen hörte die Interviewerin, deren Video besprochen wurde, zu und beteiligte sich ganz bewusst nicht an den diskutierten Erwägungen. Beim zweiten Interview, das zuerst einer diskussionslosen Einzeleinschätzung unterzogen wurde, zeigte das Spektrum der Werte, wie gut die Ratenden schon aufeinander abgestimmt waren. Es galt ja, zwei Kompetenzen weiterzuentwickeln, die Kompetenz der Interviewführung und die Kompetenz der Einschätzung.

«Andererseits: Unterschiedliche Patient/innen, unterschiedliche Interviewführung und unterschiedliche Rater/innen, so folgerten wir, generieren wohl kaum je eine erwünschte gänzliche übereinstimmende Einschätzung. Und je mehr Rater/innen an wenigen Interviews ihre Einschätzung vornehmen, desto breiter wird das Spektrum der erzielten Werte sein. Je weniger Rater/innen sich an vielen Interviews diagnostisch versuchen, umso kohärenter wird ihre Übereinstimmung in der Einschätzung ausfallen.

Am ehesten gelang uns mit der Zeit eine gute Annäherung in der Einschätzung der psychischen Struktur der Interviewten, d.h. der Wahrnehmung, der Steuerung, der Kommunikation und der Bindung in ihren jeweils weiter differenzierten Ausprägungen. Am wenigsten gelang uns eine Annäherung in der Konfliktachse, insbesondere beim ödipalen Konflikt.

Nach den Workshops verfasste ich jeweils eine inhaltliche Zusammenfassung, die an alle Rater/innen ging. Ich wollte auch die Abwesenden an den Fragen und Erkenntnissen teilhaben lassen.

Dabei motivierten mich die geführten breiten Diskussionen zunehmend, eine diagnostische Gesamtschau zum geschilderten Leiden der interviewten Patient/innen zu versuchen. Eine Gesamtschau entlang der Fragen: Woran leidet dieser Mensch? Wie zeigt sich seine Not? Wie geht er mit seinen Schwierigkeiten um? Was sind die Folgen? Meine Sätze in Prosa und nicht in Werten stiessen zunehmend auf weiterführende Resonanz bei einigen Kolleg/innen. Das Nach- und gemeinsame Weiterdenken im Mailkontakt war mir stets ein Gewinn» (ebd., S. 56f.).

Zusammenfassend kann festgehalten werden: Ein intervisorisches Voneinander-Lernen, längere, klärende, ressourcenorientierte Reflexionen, gemeinsam würdigendes Anerkennen von Unterschiedlichkeiten, Versuche einer Gesamtschau in Prosa inklusive Resonanzinteresse und weiterem Nachdenken entsprechen eher einem narrativen Zugang zu konstruktivistischen Verstehenskonzepten. Eine Steigerung der an sich erwünschten Interraterreliabilität war somit kaum mehr zu erwarten. Vielmehr hatten wie bei den Teilnehmenden in narrativen Interviews die Rater/innen eigene Perspektiven entwickelt, subjektive Sinnzusammenhänge durch kommunikative Interaktionen generiert und eigene Bedeutungsstrukturen im subjektiven Verstehen der eigenen Deutung erkennen können.

Je mehr erzählende Offenheit auch in den Einschätzungsbemühungen der Rater/innen zugelassen wird, desto geringer fällt die zu erzielende Reliabilität aus. Denn die Beziehungsgestaltung lenkt den Erzählprozess – ob im Verhältnis Patient/in und Diagnostiker/in oder im Verhältnis Diagnostiker/in und den beratenden Kolleg/inn/en. Nicht nur die zurückhaltende, respektvolle Beziehungsgestaltung zum Beispiel in narrativen Interviews bzw. dem eröffnenden Teil des OPD-Interviews, nein, auch die Beziehungsgestaltung in den Einschätzungsbemühungen der Rater/innen lenkt den Erzählprozess. Und dieser Prozess beeinflusst die Resultate. Jede Erzählung, in welchem Setting auch immer, ist ein Stück spontan-mündlicher Autobiografie. Als solche ist sie auslegungsbedürftig und verlangt einen individuellen verstehenden Zugang. Diese Auslegung darf nicht beliebig sein, soll aber auch nicht in Zahlen gegossen werden. Diese hermeneutische Arbeit ist beziehungsabhängig. Sie ist mehr Kunst als strenge Wissenschaft. Sie ist nicht zuletzt auch abhängig von der Passung zwischen den Beteiligten. Zwar weiss der Erzähler mehr über sich, auch mehr als ihm lieb ist, und mehr als er dem Zuhörenden mitteilt. Der zuhörende Interpret weiss jedoch auch oft mehr als der Erzähler. Diese unterschiedliche Passung kann in keine vergleichende Wiederholung hineinkopiert werden. Die klassischen Gütekriterien der Objektivität, der Reliabilität und der Validität von psychologischen Diagnosen können bei solchem Vorgehen nicht mehr verlangt werden. Was bleibt ist ein besonderer Erzählprozess mit der Gefahr der Beliebigkeit, aber der Chance individueller Einzigartigkeit in der unüberschaubaren Komplexität einer Lebenswelt.

Was hat der hier skizzierte Weg im diagnostischen Prozess vom Zählen zum Erzählen mit Politik zu tun? Zugespitzt: der Weg des Mathematischen und der Weg des Literarischen mit Politik? Jonas Lüscher (2017) spricht von der quantitativen Blendung in den Theorien des Zählens und von der narrativen Beliebigkeit im Erzählen. Allerdings seien Erzählungen besser geeignet, die Komplexität der Gegenwart zu beschreiben als mathematische Theorien. Das einzelne Leiden komme im Erzählen konkreter zum Vorschein als ein verrechneter Befund. Aber auch Zahlen müssten in Geschichten umgedeutet werden. Ohne Übertragung in eine Geschichte würden Zahlen nichts bedeuten. Hier schliesst auch Peter von Matt (2014) an. Die Zahlen der Naturwissenschaften

seien von den Metaphern der Geisteswissenschaften gar nicht so weit entfernt, folgert er: «Metaphern und Formeln, Modell und Erzählung: Naturwissenschaften und Geisteswissenschaften trennt kein tiefer Graben der Erkenntnis. Sie sind aufeinander angewiesen.»

Nochmals: Was haben das Zählen und das Erzählen im diagnostischen Prozess letztlich mit Politik zu tun?

Eine Psychodiagnose isoliert den Menschen zu einem Individuum. Je mehr die Befunde aufgrund von verrechenbaren Aussagen erhoben werden, umso mehr sind sie die Folge einer Politik unseres westlichen Zeitgeistes: Der Mensch als auszählbares Einzelwesen.

Zuerst die Diagnose, dann die Therapie. Was für die somatische Medizin als wissenschaftlich unerlässlich gilt, ist für eine bio-psycho-soziale Herangehensweise an den Menschen reduktionistisch. Erst die Sprache macht ja den Menschen zum Menschen. Und die Sprache ist ein Beziehungsgeschehen. Verständigung ist das Ziel. Das gelingt nur in einem längeren Prozess. Und der Verständigungsprozess kann auch eine Therapie sein, an deren Ende – nach langem Dialog, nach langer Erzählung – die Sprechenden und Zuhörenden näher an eine Diagnose gelangen, die Unterschiede kennt und Unterschiede macht. Am Anfang stehen stets Hypothesen und keine Diagnosen. Der Mensch ist primär ein soziales Wesen, die Verständigung unter den Menschen verändert sie. Und die Verständigungsprozesse beginnen mit Erzählen und nicht mit Zählen und Verrechnen. Nur: Sowohl das Zählen als auch das Erzählen sind politisch. Das Erzählen erzeugt jedoch mehr Resonanz. Wie eine Politik mit bedeutungsvollen Metaphern gegenüber einer Politik mit bedeutungsleeren Zahlen. Das Narrative und vielleicht auch die narrative Diagnostik kehren nach neuester Tendenz in die Politik zurück. Gute, überzeugende Erzählungen sind gefragt für die Meinungsbildung, nicht bloss quantifizierte Daten, Listen und Statistiken.

Politisches Handeln, so ein Merksatz von Bernauer (2015), könne charakterisiert werden als «soziales Handeln, das auf Entscheidungen und Steuerungsmechanismen ausgerichtet ist, die allgemein verbindlich sind und das Zusammenleben von Menschen regeln.»

Auch die psychologischen Diagnostiken führen zu einem sozialen Handeln, das auf Entscheidungen ausgerichtet ist, um das Zusammenleben von Menschen zu regeln.

Psychologische Diagnostik in der Krankenbehandlung muss sich sicher auf die verlangten Gütekriterien Objektivität, Reliabilität und Validität stützen. Was objektiv messbar ist, scheint mehr Einfluss auf die Gestaltung und Durchsetzung von Forderungen und Zielen in der Öffentlichkeit zu haben. Und wird so zur Macht.

Eine psychologische Diagnostik als vermeintlich objektiver Zahlenwert neigt dazu, der politischen Macht und ihrem ökonomiefokussierten Zeitgeist zu dienen. Denn seelisches Leiden muss zwingend zuerst einen vermeintlich objektiv erhobenen Krankheitswert erreichen. Darum wachsen auch die Diagnosekompendien.

Psychologische Diagnostik als Erzählung mit akzentuierter Beziehungsgestaltung bewegt sich eher in Richtung Partizipieren statt in Richtung Verschreiben, bevorzugt

Verständigung statt Herrschaft. Sie problematisiert autoritatives Expertenwissen mehr und reduziert die unübersichtliche Komplexität eines verstehbaren und auch unverständlichen Seelenlebens weniger. Zwei Wege zeigen sich: Der eine von einer letztlich in Zahlen festgehaltenen, bedeutungsarmen Diagnostik zur störungsbezogenen Therapie. Der andere von der störungsbelasteten Erzählung zur Erkenntnis, worum es den Leidenden geht.

Fazit: Eine quantitativ orientierte psychologische Diagnostik entscheidet vor allem über den Zugang zu den Versicherungsträgern und steht dadurch notwendigerweise im Dienste politischer Entscheidungen. Eine verstehensorientierte Diagnostik unterscheidet ebenso, zögert jedoch Entscheidungen hinaus. Sie steht mehr im Dienst eines erzählenden Erkennens, das die Unterschiedlichkeit des Individuellen und die Vielfalt der Möglichkeiten eigenen Lebens und Leidens anerkennt.

## Literatur

- Arbeitskreis OPD (Hrsg.). (2006). *Operationalisierte psychodynamische Diagnostik OPD-2*. Bern: Huber.
- Bernauer, T., Jahn, D., Kuhn, P. & Walter, S. (2015). *Einführung in die Politikwissenschaft*. Baden-Baden: Nomos.
- Küsters, Y. (2009). *Narrative Interviews. Grundlagen und Anwendungen* (2. Auflage). Wiesbaden: Hagener Studententexte zur Soziologie, Springer VS Verlag.
- Lienert, G.A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Auflage). Weinheim: BeltzPVU.
- Lüscher, J. (2017). Sternstunde Philosophie 06.04.2017. Schweizer Fernsehen.
- Müller-Locher, P. (2016). Der OPD-Ratingprozess. Persönliche Erfahrungen und Einsichten. In A. von Wyl, V. Tschuschke, A. Cramer, M. Koemeda-Lutz & P. Schulthess (Hrsg.), *Was wirkt in der Psychotherapie? Ergebnisse der Praxisstudie ambulante Psychotherapie zu 10 unterschiedlichen Verfahren* (S. 51–63). Gießen: Psychosozial-Verlag.
- Matt, P. von (2014). Wie wir die Welt im Geist ordnen. Vortrag zit. nach *Tages Anzeiger* vom 09.08.2017.
- Wyl, A. von, Tschuschke, V., Cramer, A., Koemeda-Lutz, M. & Schulthess, P. (Hrsg.). (2016). *Was wirkt in der Psychotherapie? Ergebnisse der Praxisstudie ambulante Psychotherapie zu 10 unterschiedlichen Verfahren*. Gießen: Psychosozial-Verlag.

## Telling stories instead of enumerating facts

The author demonstrates how the efforts towards greater reliability in diagnostic interviews run the risk of sacrificing their characteristically unstructured narrative in favour of a standardized exploration of the patient, thus, following the zeitgeist of enumerating. An example of this is the experience as a training coordinator of the evaluator community in a research study. The focus here is on the method of Operationalized Psychodynamic Diagnosis (OPD-2), that aimed to achieve a high degree of estimation consistency between the different diagnosticians.

**Keywords:** Reliability, Managing Relationships, Mathematics, Literature, Politics

### **Raccontare invece di contare soltanto**

*Sommario:* L'autore mostra come gli sforzi verso una maggiore affidabilità nelle interviste diagnostiche corre il pericolo di sacrificare la loro qualità di racconto libero a favore di un'esplorazione standardizzata del/della paziente e quindi di seguire lo spirito del tempo riguardo il numero. Ad esempio, si serve della sua esperienza come responsabile della formazione di una società di valutazione in uno studio di ricerca. In merito, egli si concentra sulla procedura di diagnostica psicodinamica operazionalizzata (OPD-2), in cui si è cercato di raggiungere un'elevata uniformità di valutazione tra i diversi diagnostici.

*Parole chiave:* affidabilità, forma di relazione, matematica, letteratura, politica

### **Der Autor**

*Peter Müller-Locher*, Dr. phil., Msc in Organisationsentwicklung, ist niedergelassener psychologischer Psychotherapeut (ASP), Daseinsanalytiker, Supervisor und Gruppenanalytiker. Er ist Vorsitzender der Kommission für Qualitätssicherung der Schweizer Charta für Psychotherapie und Teilnehmer an der Praxisstudie ambulante Psychotherapie (PAP-S) in verschiedenen Chargen.

### **Kontakt**

[peter.mueller-locher@bluewin.ch](mailto:peter.mueller-locher@bluewin.ch)