

12. Vorlesung: Gütekriterien und Qualitätssicherung qualitativer Methoden

Meine Damen und Herren!

In der letzten Vorlesung gehe ich auf Gütekriterien und Qualitätssicherung qualitativer Methoden ein. Dies ist sicher auch für die praktische Diagnostik von größter Bedeutung, die Konzepte stammen aber weitgehend aus der Forschung. Qualitätssicherung ist für die qualitative Forschung – der ja häufig vorgeworfen wird, ihre Ergebnisse nicht wissenschaftlich ausweisen zu können – von großer Bedeutung.

Als erstes werde ich kurz auf unterschiedliche Wahrheitstheorien eingehen, anschließend geht es um die Vermittlungsschritte und Fehlerquellen bei der Datenerhebung und Dateninterpretation. Danach gehe ich ein auf die Übertragbarkeit traditioneller psychometrischer Gütekriterien auf qualitative Methoden wie Interviews oder Beobachtungsmethoden. Zum Schluss möchte ich Ihnen zeigen, wie Konzepte der Qualitätssicherung für die qualitative Diagnostik und Forschung nutzbar gemacht werden können.

1 Wahrheitstheorien

Wahrheitstheorien sind ein Thema der Erkenntnis- oder Wissenschaftstheorie.

Stichworte Wahrheitstheorien

- Abbildtheorie
Theorien sind wahre Abbilder der Realität
- Kritischer Rationalismus
Theorien haben hypothetische Gültigkeit, so lange sie nicht falsifiziert sind
- Konsenstheorie
Wahrheit wird festgestellt durch den Konsens von Experten
- Konstruktivismus
Wahrheit wird in einem sozialen Prozess konstruiert

Nach der *Abbildtheorie* sind Theorien einfach Abbilder der Realität, es gibt eine Eins-zu-Eins-Übereinstimmung gewisser Eigenschaften der Realität mit der Theorie. Dies ist die Position des naiven Realismus, der davon ausgeht, eine Theorie ist ein Abbild der Realität.

Im *Kritischen Rationalismus*, wie er von *Karl Popper* in seinem Klassiker "Logik der Forschung" vertreten wird, tritt der Abbildcharakter in den Hintergrund. Theorien sind nach Popper so lange gültig oder wahr, wie sie nicht falsifiziert worden sind, d.h. wie es keine Beweise oder Experimente für ihre Ungültigkeit gibt. Theorien haben damit hypothetischen Charakter, sie müssen falsifiziert werden können. Der Prozess der Erkenntnisgewinnung wird als ein kumulativer Prozess vorgestellt, wobei die Wissenschaft immer weiter fortschreitet und die Erkenntnisse immer wahrer, immer zutreffender werden durch Ablehnung falscher Theorien und Aufstellung neuer Theorien.

Der *Kritische Rationalismus* lässt sich nach den Untersuchungen des Wissenschaftshistorikers *Thomas Kuhn* nicht mehr halten. *Kuhn* hat festgestellt, dass die Forschung von *Paradigmen*

abhängig ist, die die Sichtweise der Wissenschaftler auf ihren Gegenstand bestimmen. Im Zuge der wissenschaftlichen Entwicklung finden wissenschaftliche Revolutionen statt, in denen ein Paradigma (eine Sichtweise auf die Welt) das nächste ablöst. Diese Paradigmen können nicht mehr mit dem Verwerfen von Hypothesen erklärt werden, sondern es ist nötig, für diese Konzeption auf die Gemeinschaft der Forscher, auf die Gemeinschaft der Wissenschaftler zurück zu greifen, die eine gemeinsame Weltsicht teilen. Damit kommen wir zur Konsenstheorie der Wahrheit.

Die *Konsenstheorie* besagt, dass Wahrheit durch den Konsens von Experten festgestellt wird. Wenn Sie an das Gutachterwesen denken – etwa in der Politikberatung oder in anderen Bereichen – kann man sagen, in unserem Alltag ist die Konsenstheorie weitgehend Praxis. Wir gehen etwa davon aus, dass der Konsens ausgewiesener Experten uns sagt, ob es notwendig ist, 400.000 Kühe wegen der Gefahr der Infektion mit BSE zu schlachten oder nicht, um nur ein aktuelles Beispiel zu nehmen. *Habermas* ist ein Anhänger der Konsenstheorie. Er vertritt die These, dass durch einen herrschaftsfreien Diskurs und durch die Geltungsansprüche der Alltagssprache, wenn sie nur eingehalten werden, am ehesten so etwas wie eine Wahrheitsfindung möglich ist.

Die *konstruktivistische Wahrheitstheorie* geht einen Schritt weiter. Tatsächlich zeigt auch die Praxis heutzutage, wenn ein Gutachten durch ein Gegengutachten mit einer völlig anderen Position konterkariert wird, dass der Konsens der Experten nicht zu *einer* alleingültigen Wahrheit führt, sondern dass wissenschaftliche Kontroversen bestehen bleiben. Beim Konstruktivismus wird die Konzeption *einer* Wahrheit ganz aufgegeben, man geht vielmehr davon aus, dass Wahrheit in einem sozialen Prozess konstruiert wird und dass dieser Prozess je nach Perspektive zu ganz unterschiedlichen Wahrheiten führen kann. Es erhebt sich beim radikalen Konstruktivismus die Frage, wie weit diese Auffassung nicht zum Relativismus, zur Beliebigkeit führt, zu dem, was der Wissenschaftstheoretiker *Feyerabend* "anything goes" genannt hat. Im Sozialkonstruktivismus gibt es aber noch einen Anker zur Bewährung von Theorien: Man nimmt nicht mehr an, dass die eine Theorie richtig ist, die andere falsch, sondern dass es sich um unterschiedliche Konstruktionen handelt, wobei entscheidend ist, wie weit eine Konstruktion für eine konkrete Fragestellung zur Lösung eines bestimmten Problems besser geeignet ist als die andere.

Zwischen der Konsenstheorie und der konstruktivistischen Wahrheitstheorie gibt es fließende Übergänge, trotzdem werden unterschiedliche Akzente gesetzt. In der qualitativen Forschung, in der wir auf die Subjektivität des Forschers und der Beforschten abheben und diese in den Mittelpunkt stellen, scheint mir – und ich befinde mich da in Übereinstimmung mit den meisten Vertretern dieser Forschungsrichtung – ein gemäßiger Konstruktivismus die angemessenste wissenschaftstheoretische Position, die es erlaubt, den Forschungsprozess als soziale Konstruktion zu reflektieren.

2 Vermittlungsschritte und Fehlerquellen bei der Datenerhebung

2.1 Konstruktionen im Forschungsprozess

Betrachten Sie folgendes Schema zu "Konstruktionen im Forschungsprozess":

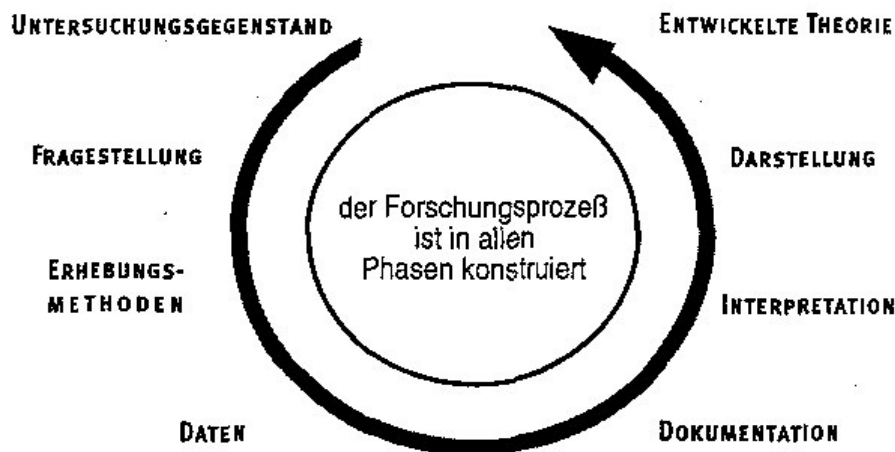
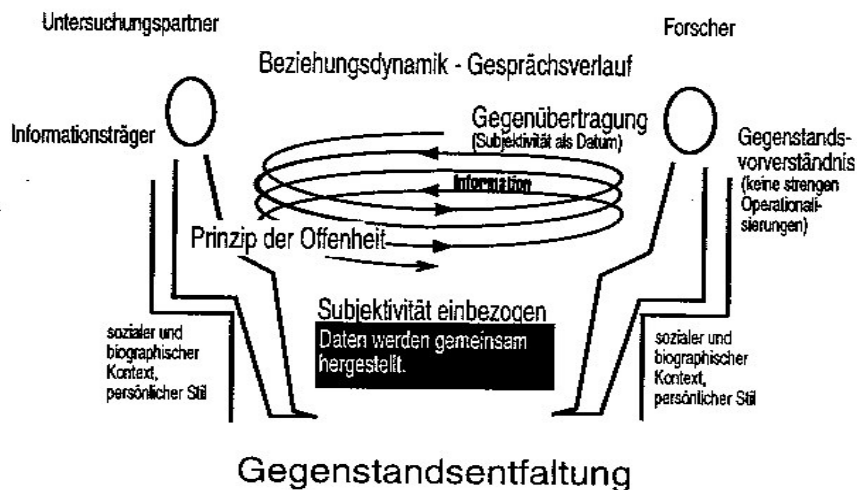


Abb.: Konstruktionen im Forschungsprozess (Steinke 1999)

Wie Sie sehen, wird der komplette Forschungsprozess dargestellt: von der Auswahl und Definition des Untersuchungsgegenstandes über die Entwicklung der Fragestellung, die Auswahl und Durchführung der Erhebungsmethoden, die Erstellung der Daten, die Dokumentation, die Interpretation, die Darstellung der Forschungsergebnisse bis hin zur entwickelten Theorie. Alle diese Schritte lassen sich auffassen als soziale Prozesse der Konstruktion von Wirklichkeit, in denen je nach Auswahl der Daten, Methoden, usw. *unterschiedliche Sichtweisen der Welt* konstruiert werden können.

Wenn Sie qualitative oder quantitative Methoden auf den gleichen Untersuchungsgegenstand – z. B. Armut in der Großstadt – anwenden, erhalten Sie unterschiedliche Modellkonstruktionen der Wirklichkeit. Es ist nicht sinnvoll zu sagen, das eine Modell ist richtig, das andere ist falsch, sondern es geht um die Angemessenheit der jeweiligen Methode zur Lösung von Problemen.

Wir wollen uns die Konstruktionen im Prozess der Datenerhebung am Beispiel eines Interviews einmal genauer ansehen:



Gegenstandsentsfaltung durch:

- Reflexiven Charakter der Forschung
- Gestaltung der Erhebungssituationen (alltagsorientiert, nach dem Prinzip der Offenheit, Zulassen von Gegenübertragungen)
- Theoretical Sampling
- Methodenwahl
- den/die Rezipienten

Abb.: Gegenstandsentsfaltung im Interview (Steinke 1999)

Hier finden wir beim Interviewten (dem Informationsträger): seine soziale und biografische Situation und Interessenlage, der Kontext des Interviews, der persönliche Stil des Befragten u.s.w. Beim Interviewer (Diagnostiker oder Forscher) finden sich die gleichen Aspekte - hinzu kommen seine fachliche Kompetenz, sein Gegenstandsvorverständnis und die gewählte Interviewmethode. All das wird die Art seiner Fragestellung, die Art der einzelnen Fragen, die Art der Nachfragen bestimmen. Das Gegenstandsvorverständnis lässt sich nicht streng operationalisieren, weil es im Alltagswissen verankert ist.

Entscheidend ist, dass der Gegenstand bzw. die über ihn ausgetauschten Informationen gemeinsam entfaltet, gemeinsam entwickelt werden. Jedes Interview stellt eine Konstruktion von zwei Individuen, zwei Subjekten dar. Es geht hier nicht um eine Beziehung zwischen Untersuchungsobjekt und Untersuchungssubjekt, sondern zwischen selbstständig reflexionsfähigen Subjekten, die den Gegenstand entfalten.

Sie erinnern sich, ich hatte früher das zentrale Konzept des Arbeitsbündnisses eingeführt. Es ist ein roter Faden durch diese Veranstaltung: Immer kommt es auf die Freiwilligkeit des Interviewten, auf die Einwilligung, das Sich-Einlassen in die Forschungssituation an, auf die Vermittlung der Fragestellung an den Interviewten, damit er oder sie in entsprechender Weise kooperieren kann.

3 Gütekriterien

3.1 Die klassischen Test-Gütekriterien

Stichworte Test-Gütekriterien

- **Objektivität**
Grad, in dem ein diagnostisches Ergebnis unabhängig vom Untersucher ist
 - Durchführungsobjektivität
 - Auswertungsobjektivität
 - Interpretationsobjektivität
- **Reliabilität (Zuverlässigkeit)**
Grad der Genauigkeit bzw. Reproduzierbarkeit eines diagnostischen Ergebnisses
 - Paralleltest-Reliabilität
 - Retest-Reliabilität
 - innere Konsistenz
- **Validität (Gültigkeit)**
Grad der Genauigkeit, mit der ein diagnostisches Ergebnis das Merkmal, das es erfassen soll, tatsächlich erfasst, wobei zwischen interner und externer Reliabilität unterschieden wird
 - inhaltliche Validität
 - Konstruktvalidität
 - Kriteriumsvalidität (Übereinstimmung – Prognose)
 - ökologische Validität

Beim Einsatz quantitativer Methoden (z.B. Tests oder Fragebogen) werden zur Beurteilung der Güte oder Qualität der Datenerhebung drei so genannte *Haupt-Gütekriterien* unterschieden: *Objektivität, Reliabilität und Validität*.

- Das Kriterium der *Objektivität* gibt an, wie weit das Ergebnis eines Tests abhängig davon ist, wer die Untersuchung durchgeführt hat, wobei zwischen der Durchführungsobjektivität, Auswertungsobjektivität und Interpretationsobjektivität unterschieden wird.
- Bei der *Reliabilität* geht es um die Wiederholbarkeit und damit Zuverlässigkeit der Datenerhebung in Anlehnung an das Messkonzept in der Physik, wobei man davon ausgeht, dass die zu "messenden" Merkmale über die Zeit hinweg stabil sind.
- Das praktisch wichtigste Gütekriterium ist die *Validität* oder *Gültigkeit*. Die Validität erfasst die Genauigkeit, mit der von den erhobenen Daten auf das Merkmal geschlossen werden kann, das erfasst werden soll (z.B. vom Ergebnis eines Fragebogens zur Erfassung der Angst auf die tatsächlich vorhandene Angst).

Es wird zwischen *interner* und *externer* Validität unterschieden, wobei die interne Validität sich auf die innere Stimmigkeit und Konsistenz des Erhebungsinstruments bezieht, während die externe Validität sich auf die Korrelation mit einem Außenkriterium bezieht (im Beispiel: ein Maß für die tatsächlich vorliegende Angst – falls es das gibt). Da die Datenerhebung gewöhnlich kein Selbstzweck ist, sondern bei jeder Datenerhebung auf Alltagsphänomene in Realsituationen geschlossen werden soll, ist die externe Validität letztendlich entscheidend für die praktische Anwendung von Forschungsergebnissen. Man unterscheidet verschiedene Arten externer Validität: *Inhaltsvalidität, Konstruktvalidität, Kriteriumsvalidität, ökologische Validität*.

Für das Ausmaß, in dem jedes dieser Gütekriterien erfüllt ist, lässt sich in quantitativen Untersuchungen bei Vornahme geeigneter Kontrolluntersuchungen ein Schätzmaß in Form eines Korrelationskoeffizienten berechnen.

Es leuchtet ein, dass diese Gütekriterien aus verschiedenen Gründen nicht umstandslos zur Beurteilung der Güte qualitativer Methoden herangezogen werden können. In der Diskussion um Gütekriterien für die qualitative Forschung wird aber von manchen Forschern versucht, die klassischen Gütekriterien zu modifizieren, so dass sie den Besonderheiten qualitativer Methoden angemessener sind.

Die Gegenposition zur Modifikation der traditionellen Gütekriterien für qualitative Methoden geht davon aus, dass die Gütekriterien quantitativer Verfahren völlig ungeeignet für das Feld qualitativer Forschung seien, in dem neue und eigenständige Güte- bzw. Qualitätskriterien zu entwickeln seien.

Für die Praxis angewandter Sozialforschung möchte ich Ihnen – jenseits der zum Teil ideologisch geführten wissenschaftstheoretischen Kontroversen – empfehlen, den logischen Grundgedanken der klassischen Gütekriterien sinngemäß auf qualitative Methoden, z.B. auf Interviews, anzuwenden, aber darüber hinaus auch eigenständige *prozessorientierte Qualitätskriterien* für qualitative Untersuchungen heranzuziehen. Darum soll es in den nächsten Abschnitten gehen.

3.2 Gütekriterien im Vergleich

In folgender Tabelle werden die traditionellen Gütekriterien der quantitativen Forschung aus statistischen (Test-)Verfahren ihren qualitativen "Verwandten" gegenüber gestellt, also Gütekriterien die der Logik qualitativer Verfahren eher angemessen sind

Gütekriterien im Vergleich	
FÜR TESTS <i>(ergebnisbezogen in Form eines Koeffizienten)</i>	FÜR QUALITATIVE METHODEN <i>(bezogen auf Schlussfolgerung als argumentative Begründung)</i>
Methodeninterne Gütekriterien	
Objektivität	Intersubjektivität
Reliabilität	Reproduzierbarkeit
interne Validität	interne Validität
Externe Gütekriterien (externe Validität)	
Inhaltsvalidität	Inhaltsvalidität
Konstruktvalidität	Konstruktvalidität
Kriteriumsvalidität	Kriteriumsvalidität
Ökologische Validität	Ökologische Validität

Objektivität ist ein zentrales methodeninternes Gütekriterien traditioneller Forschung. Objektivität strebt immer eine Standardisierung der Situation an, die aber in der qualitativen Forschung aus guten Gründen abgelehnt wird.

In der qualitativen Forschung spielt die Subjektivität von Forscher und Beforschten eine wichtige Rolle im Erkenntnisprozess und kann nicht eliminiert werden. Dies konnten Sie oben am Beispiel der sozialen Interaktion zwischen Interviewer und Interviewtem in Verlauf eines qualitativen Interviews sehen. Die systematische Reflexion der Subjektivität der beteiligten

Menschen und ihrer Deutungen sozialer Wirklichkeit trägt zur Qualität und Nachvollziehbarkeit qualitativer Forschungsergebnisse bei.

Als logischer Kern des klassischen Objektivitätskriteriums kann hier auf den qualitativen Forschungsprozess die intersubjektive Übereinstimmung verschiedener Untersucher in der Analyse desselben Forschungsgegenstandes übertragen werden. Stellen Sie sich vor, dass in einer qualitativen Untersuchung Forschungsinterviews von Interviewer X und Interviewerin Y geführt werden. Selbst wenn die Interviewverläufe ganz unterschiedlich sind, wäre zu fordern, dass die Forscher letzten Endes zu den gleichen Ergebnissen kommen. Wenn es am Ende der qualitativen Analysen zu Abweichungen in den Ergebnissen kommt, müssen diese begründet werden, denn es liegt dann ein Fall von nicht vorhandener Übereinstimmung vor. Die Objektivität der traditionellen Testdiagnostik findet danach ihre qualitative Entsprechung in der intersubjektiven Übereinstimmung, in der *Intersubjektivität* der Schlussfolgerungen und Ergebnisse einer Forschungsgruppe.

Die *Methoden*, diese Intersubjektivität qualitativer Ergebnisse zu gewährleisten und zu kontrollieren, unterscheiden sich von den Prüfverfahren quantitativer Ansätze. Statt Prozentwerte der Übereinstimmung zu berechnen, wird ein kommunikativer Reflexionsprozess in Gang gebracht, ein Diskurs über die unterschiedlichen Sichtweisen – etwa in einer Forschungs-supervision - um zu einem Konsens in der Forschergemeinschaft zu kommen.

Der rationale, logische Kern des klassischen Kriteriums der *Reliabilität* ist die Forderung nach Reproduzierbarkeit der Ergebnisse, die auch auf qualitative Forschungsgegenstände anzuwenden ist. Wenn es sich um einen Zusammenhang handelt, der sich mit der Zeit nicht grundlegend geändert hat (z.B. eine zentrale Persönlichkeitseigenschaft), der also einigermaßen zeitstabil ist, dann sollte bei einer zweiten qualitativen Untersuchung in der Diagnostik im Prinzip das gleiche Ergebnis erzielt werden.

Die Reproduzierbarkeit als Kern der Reliabilität muss allerdings als ein relatives Kriterium angesehen werden. Selbst bei klassischen Testverfahren sind keine statischen Bedingungen gegeben. Denn wir alle sind historische Wesen – um es mit einem antiken Philosophen zu sagen: "Du steigst nie zwei Mal in den gleichen Fluss". Die Wiederholbarkeit eines Tests ist im Grunde eine Fiktion und bei der wiederholten Durchführung eines Interviews kann ich trotz desselben Kontextes und derselben Fragestellung nicht davon ausgehen, dass der Interviewverlauf derselbe ist.

Methoden, um die Reliabilität – beispielsweise in der Forschungsarbeit mit qualitativen Interviews – zu fördern, sind u.a. Schulungen von Forschungspersonal für die Erhebung und Auswertung von Forschungsinterviews. So können die Daten unter vergleichbaren Rahmenbedingungen (Eingangsfrage, Leitfadenverwendung, Fragestil u.a.) erhoben und eine gleichartige Kodierung des Datenmaterials auch bei wechselnden Kodierern (*Interkoderreliabilität*) gewährleistet werden.

3.3 Interne Validierung von Interviews

Interne Validierung von Interviews

- *Strukturelle Voraussetzungen*
 - Kommunikative Kompetenzen
 - Freiwilligkeit
 - Interessenklärung
- *Rollen und Motive*
 - Rollenverteilung
 - Kommunikationsstil
 - Soziale Erwünschtheit
- *Qualität der Beziehung*
 - Vertrauen
 - Krisen
- *Verständlichkeit*
 - Sprache und Lebenswelt
 - Verständnissicherung
- *Angemessenheit*
 - Arbeitsbündnis
- *Wahrheit*
 - Widersprüche
 - Externe Hinweise
- *Aufrichtigkeit*
 - Diskrepanzen (nonverbale Hinweise?)
- *Hinweise auf Lücken und Verzerrungen*

Auf der Grundlage der *Theorie des kommunikativen Handelns* von Jürgen Habermas (1988) können Interviews bezüglich ihrer internen Stimmigkeit oder internen Konsistenz geprüft werden. Habermas formuliert die Ihnen aus früheren Vorlesungen vertrauten *Geltungsansprüche* für kommunikatives Handeln, die zugleich als Prüfkriterien für die Beurteilung von kommunikative Äußerungen (in Gesprächen, Interviews und allgemein in jeder Art sprachlicher Daten) dienen können:

- **Verständlichkeit:** Wie weit haben die Gesprächspartner ihre jeweiligen Äußerungen verstanden? Welche Rolle spielten dabei Sprache und Lebenswelt der Beteiligten (Gemeinsamkeiten und Unterschiede der soziokulturellen Codes)? Wie weit hat der Interviewer bei Unklarheiten durch seine Nachfragen zur Klärung beigetragen?
- **Angemessenheit:** Wie weit wurde das Arbeitsbündnis für das Forschungsgespräch (Interessen und Rollen der Gesprächspartner) vermittelt und tatsächlich vollzogen? Das heißt, wie weit ist der Interviewte im Sinne der Forschungsfrage kooperativ gewesen?
- **Wahrheit:** Wie weit bestehen interne Widersprüche in den Aussagen bzw. Widersprüche zu bekannten externen Fakten?
- **Aufrichtigkeit:** Wie weit wurden die subjektiven Gefühle, Absichten und Wünsche aufrichtig geäußert? Gibt es Hinweise auf Lücken, Verzerrungen, Verschwiegenes und Verdrängtes? Gibt es Diskrepanzen zwischen verschiedenen Kommunikationsebenen (Körpersprache, verbale Aussagen)?

Bei der internen Validierung von Interviews überprüfen Sie die Vertrauenswürdigkeit und Glaubwürdigkeit der Aussagen Ihrer Interviewpartner. Dafür schauen Sie sich die Kommunikationssituation an und prüfen zuerst einmal die strukturellen Voraussetzungen des Interviewten:

Wie weit ist er/sie in der Lage, sich zu dem Forschungsthema zu äußern? Wie sind seine/ihre kommunikativen Kompetenzen? Er/Sie kann beispielsweise in seiner kognitiven Leistungsfähigkeit durch Drogen beeinträchtigt sein.

Er/Sie kann aufgrund eines Schlaganfalls einen Sprachdefekt haben, sie/er kann moralische Hemmungen haben über ein Thema offen zu sprechen usw. Wichtig ist dabei die Interessenklärung und Motivation zur Gesprächsteilnahme der befragten Personen; wenn beispielsweise keine Freiwilligkeit in der Gesprächsanbahnung gewährleistet ist, sondern sie/er zur Teilnahme zwangsverpflichtet wird, können Sie nicht erwarten, dass ein verständigungsorientiertes Handeln zustande kommt.

Wenn der Interviewte bestimmte Interessen mit dem Gespräch verbindet wie z.B. bei einem Interview in einer Bewerbungssituation oder bei einem polizeilichen Verhör, müssen unter Umständen an der Erfüllung der *Geltungsansprüche* (Verständlichkeit, Wahrheit, Wahrhaftigkeit und Angemessenheit) Abstriche gemacht werden, da hier vom Interviewten vor allem Äußerungen entsprechend der unterstellten sozialen Erwünschtheit zu erwarten sind.

Außerdem sollte die Qualität der Beziehung zwischen Interviewer und Interviewtem eingeschätzt werden: Wie weit ist Vertrauen vorhanden, so dass auch persönlichere Dinge im Interviewkontext angesprochen werden können? Gab es möglicherweise Gesprächskrisen, die unter Umständen von einem bestimmten Punkt an das Vertrauen gemindert haben? Welche Auswirkungen hatte dies auf die Äußerungen des Interviewten?

3.4 Externe Validität

Wenden wir nun den Blick von den erhobenen Daten und gewonnenen Forschungsergebnissen wiederum auf die untersuchten Phänomene des sozialen Lebens, so kommen wir zum Gütekriterium der externen Validität. Hierbei geht es um die Realitätshaltigkeit und Übertragbarkeit der Forschungsergebnisse.

Stichworte Externe Validität

Externe Validität liegt vor, wenn das in einer Untersuchung gefundene Ergebnis zu Phänomen A auch jenseits der Untersuchungssituation Aussagen über A erlaubt.

- Für quantitative und qualitative Untersuchungen gleichermaßen bedeutsam
- Vergleich zwischen Untersuchungsergebnis und Außenkriterium
- Externe Validität qualitativer Methoden durch systematische Einzelfalluntersuchungen

Nehmen wir an, ein psychologischer Diagnostiker kommt aufgrund eines Interviews zu dem Ergebnis, ein Polizeibeamter sei für den Streifendienst nicht mehr einsatzfähig aufgrund seiner Angst, die er nach einer schweren Verletzung auf der Streife entwickelt hat. Das relevante Außenkriterium, das in beruflichen Situationen dieses Polizisten festgestellt werden kann, ist die wahrnehmbare große Angst, die einen weiteren Einsatz im Streifendienst für diesen Polizeibeamten als nicht zumutbar erscheinen lässt, da er in diesem Zustand eine Gefährdung für sich und andere darstellen könnte. Der Schluss vom Ergebnis auf die Alltagssituation unterscheidet sich in einer qualitativen Untersuchung formal nicht von einer Vorhersage, die aufgrund eines Tests in der psychometrischen Diagnostik gemacht werden kann.

Die verschiedenen *Aspekte der externen Validität* lassen sich in ihren zentralen Sinngehalten gut auf qualitative Forschung übertragen.

Von *Inhaltsvalidität* können wir auch bei einem qualitativen Forschungsinterview sprechen. Wenn es beispielsweise um Konflikte in einer Partnerschaft geht, wäre ein Interview dann inhaltsvalide, wenn genau diese Konflikte auch im Interview zum Thema werden.

Bei der *Konstruktvalidität* geht es um die Frage, wieweit ein Interview etwa ein psychoanalytisches Konstrukt wie z. B. Widerstand wirklich erfasst. In der Operationalisierten Psychodynamischen Diagnostik (OPD) geht z.B. es darum, psychoanalytische Konstrukte (innerer Konflikt, Bindungsverhalten, Beziehungsverhalten, Art der Beziehungsaufnahme) anhand der inhaltlichen Ergebnisse eines diagnostischen Interviews *und* anhand der Interaktion im Interview konstruktvalide erfassen zu können.

Bei der *Kriteriumsvalidität* geht es um das Ausmaß der Übereinstimmung der Ergebnisse mit einem Außenkriterium.

Die *ökologische Validität* bezieht sich auf die Ähnlichkeit der Kontextbedingungen zwischen Untersuchung und Alltagsrealität als Voraussetzung für die Übertragbarkeit der Forschungsergebnisse auf die Lebenswelt der Beforschten.

4 Qualitätssicherung qualitativer Methoden

Das Beispiel der Verzerrungsmöglichkeiten bei biografischen Interviews hat deutlich gemacht, dass ein angemessenes Konzept zur Beurteilung der Qualität qualitativer Methoden sich nicht allein auf die Ergebnisse einer Untersuchung beschränken darf, sondern den gesamten Prozess von der Planung über die Datenerhebung bis zur Interpretation umfassen muss. Ein Ansatz zur Sicherstellung und Beurteilung der *Qualität von Prozessen*, der in letzter Zeit auch auf die Qualität von Diagnostik und Forschung übertragen wird, stammt aus der Industrie und Wirtschaft.

Als Faktor des Wettbewerbsvorteils haben amerikanische und japanische Unternehmen entdeckt, dass es wichtig ist, systematisch die Qualität ihrer Produkte oder Dienstleistungen zu untersuchen. In den letzten Jahren wird in allen möglichen Bereichen die Forderung erhoben, Verfahren der Qualitätssicherung bzw. des Qualitätsmanagements und im Zusammenhang damit auch eine Qualitäts-Zertifizierung für Dienstleistungen einzuführen. Das gilt z.B. so unterschiedliche Dienstleistungen wie die Lehre an den Universitäten, die Pflegeleistungen in der Medizin und die Psychotherapie. Ich nehme an, dass Sie alle mit dieser *Qualitätsbewegung* schon konfrontiert worden sind.

Nach Definition der Industrienorm DIN oder ISO 8402 wird Qualität definiert als "Gesamtheit von Merkmalen (und Merkmalswerten) einer Einheit [*Produkt oder Dienstleistung*] bezüglich ihrer Eignung, festgelegte und vorausgesetzte Erfordernisse zu erfüllen".

Die einzelnen Merkmalswerte eines Produkts oder einer Dienstleistung können dabei unterschiedliche Werte von *sehr schlecht* bis *sehr gut* annehmen.

Solche Definitionen sind sehr abstrakt. Wichtig für uns sind die unterschiedlichen Dimensionen bzw. Aspekte von Qualität. Eine auch für wissenschaftliche Dienstleistungen anwendbare Systematik unterscheidet zwischen den folgenden vier Qualitätsaspekten:

1. konzeptionelle Qualität
2. Strukturqualität
3. Prozessqualität
4. Ergebnisqualität

Normalerweise denkt man nur an die Ergebnisqualität - die in der Diagnostik und Forschung durch die vorher diskutierten Gütekriterien erfasst wird. Bei der prozessorientierten Qualitätssicherung geht man jedoch davon aus, dass zur Steigerung der Ergebnisqualität der gesamte Herstellungsprozess eines Produktes oder einer Dienstleistung optimiert werden muss.

Stichworte Qualitätsaspekte qualitativer Methoden

1. *Konzeptionelle Qualität*

- Fundierter theoretischer Hintergrund
- sinnvolle und eindeutige Fragestellung
- Gegenstandsangemessenheit der Methode
- Kontextangemessenheit der Methode und des Vorgehens

2. *Strukturqualität*

- Qualifikation/Training des Untersuchers
- angemessene organisatorische Rahmenbedingungen
- Qualitätssicherungsmaßnahmen (z.B. Supervision)
- Sicherstellung angemessener Partizipation

3. *Prozessqualität*

- Einhalten ethischer Standards
- Angemessene situative Bedingungen
- Professionelle Planung und Durchführung
- Selbstreflexion/Gegenübertragung
- Angemessene Rückmeldung
- Angemessene Dokumentation
- Angemessene Interpretation/Urteilsbildung

4. *Ergebnisqualität*

- Einhaltung der (qualitativen) Gütekriterien
(Intersubjektivität, Replizierbarkeit, Validität)
- Relevanz
- Umsetzung
- Evaluation

Welche praktische Bedeutung haben diese Konzepte in der Anwendung auf qualitative Forschung? Was heißt es, ein qualitatives Vorgehen oder eine Forschungsmethode besitzt *konzeptionelle Qualität*? Das Vorgehen muss theoretisch fundiert sein, es muss sich um eine sinnvolle und eindeutige Fragestellung handeln, die eingesetzten Methoden der Fragestellung müssen aber auch dem jeweiligen Kontext angemessen sein. Jemanden, der kaum lesen und schreiben kann, kann man z.B. nicht mit einem Fragebogen untersuchen. Das wäre fehlende Kontextangemessenheit.

Strukturqualität bezieht sich auf die strukturellen Voraussetzungen, die gegeben sein müssen, damit in einer Einrichtung oder einem Projekt Qualität erzielt werden kann. Zuerst einmal sollten Sie die Qualifikation des Forschers anschauen. Er/Sie muss z.B. in der Lage sein, ein Interview kompetent durchzuführen bzw. mit anderen Methoden professionell zu arbeiten. Dann müssen die organisatorischen Rahmenbedingungen gegeben sein. Es muss ein geeigneter Raum zur Verfügung stehen sowie ein funktionsfähiges Gerät für die Aufzeichnung der Daten. Dann sollten regelmäßig Qualitätssicherungsmaßnahmen wie z.B. Supervision durchgeführt werden, d.h. die Qualität sollte Thema in der Zusammenarbeit sein und es sollte verbindliche Qualitätskriterien geben, usw. Als Letztes ist eine angemessene Partizipation der Untersuchten an der Durchführung sicher zu stellen. Wir gehen in der Sozialforschung davon aus, dass wir es mit mündigen Bürgern zu tun haben, die in angemessener Weise in den Pro-

zess der Untersuchung einbezogen werden müssen. Die Aktionsforschung zur partizipativen Stadtteilentwicklung ist ein Beispiel für weitreichende Partizipation in der Forschung. In anderen Projekten gewährleistet die informierte Zustimmung ein Minimum an Partizipation.

Als nächstes kommen wir zur *Prozessqualität*, die sich auf die konkrete Durchführung der einzelnen Untersuchung bezieht. Alle Aspekte der Strukturqualität wirken sich selbstverständlich auf die einzelne Durchführung aus. Unter dem Aspekt der Prozessqualität wird geprüft, wieweit die erforderlichen Qualitätsanforderungen bei der Durchführung *im Einzelfall* erfüllt wurden. Auch ein kompetenter Interviewer (Aspekt der Strukturqualität) kann sich bei einem einzelnen Interview inkompetent verhalten haben (Aspekt der Prozessqualität des jeweiligen Falls). Bei der Qualität der Durchführung geht es zunächst um das Einhalten ethischer Standards, d.h. um informierte Zustimmung, Datenschutz und andere ethische Überlegungen. Ein eklatanter Fehler wäre etwa die Durchführung eines Stressinterviews bei jemandem, der unter Angstzuständen leidet oder ähnliche Dinge, von denen man immer wieder in den Zeitungen liest. Angemessene situative Bedingungen müssen gegeben sein. Professionelle Planung und Durchführung des Interviews, Selbstreflexion, Beachten der Gegenübertragung sind weitere wichtige Komponenten des Untersuchungsprozesses, schließlich eine angemessene Rückmeldung der Ergebnisse an den Interviewpartner entsprechend der Fragestellung und dem Kontext. Wichtig ist weiterhin eine angemessene Dokumentation, z.B. durch Bandaufzeichnung und durch ein Interviewprotokoll. Am Ende steht die Angemessenheit der Interpretation nach einem Regel-geleiteten Verfahren, das zu einer nachvollziehbaren Urteilsbildung führt. Auch die Supervision – deren generelle Verfügbarkeit ein Aspekt der Strukturqualität ist – kann beim Einzelfall fehlen oder unangemessen verlaufen sein.

Die *Ergebnisqualität* ist das Entscheidende, wozu der ganze Aufwand betrieben wird: es geht letzten Endes um die inhaltliche Qualität der Ergebnisse einer Untersuchung. Hier kommen die klassischen Gütekriterien zum Tragen, allerdings in modifizierter, für qualitative Methoden angemessener Form: Intersubjektivität, Reproduzierbarkeit und als wichtigstes Kriterium externe Validität. Ein weiterer wichtiger Aspekt ist die Relevanz des Ergebnisses: was nutzt es, wenn externe Validität für ein bestimmtes Ergebnis gegeben ist, aber das Ergebnis ist zur Lösung des jeweiligen Praxisproblems wenig relevant?

Schließlich ist unter Qualitätsgesichtspunkten zu fragen, wieweit das Ergebnis auch umgesetzt und wieweit die Umsetzung evaluiert wird.

Ich möchte Sie in diesem Zusammenhang hinweisen auf einen interessanten Vorschlag zur Beurteilung der Qualität von Daten bzw. Informationen, der aus dem *Wissensmanagement* stammt. Wissensmanagement ist ein neuer Modebegriff, darunter werden Prozesse des Wissenserwerbs und der Wissensnutzung in Organisationen und in der Organisationsentwicklung zusammengefasst.

In der folgenden Übersicht sehen Sie, welche Qualitätskriterien sich für die "Ware Information" in Wirtschaft und Industrie bewährt haben (s. Huang, Lee, Wang, 1999). Die Kriterien für *Intrinsische Qualität* korrespondieren teilweise mit den traditionellen Gütekriterien der Testdiagnostik. Diese Aspekte werden ergänzt durch zusätzliche Qualitätskriterien, die sich auf Kontext, Repräsentation und Zugänglichkeit der Informationen beziehen.

Informationsqualität im Wissensmanagement	
Kategorie	Dimensionen
Intrinsische Qualität	<ul style="list-style-type: none"> - Genauigkeit - Objektivität - Glaubwürdigkeit - Reputation

Kontextbezogene Qualität	<ul style="list-style-type: none"> - Relevanz - Mehrwert - Aktualität - Vollständigkeit - Umfang
Repräsentationsbezogene Qualität	<ul style="list-style-type: none"> - Interpretierbarkeit - Verständlichkeit - Präzision - Konsistenz
Qualität der Zugänglichkeit	<ul style="list-style-type: none"> - Zugang - Sicherheit

Abschließend möchte ich auf den Unterschied zwischen Qualitäts-Leitlinien und Kriterien (Standards) eingehen. Die dargestellten Qualitätsaspekte besitzen die Funktion von *Leitlinien*, die dem qualitativen Forscher verdeutlichen, worauf bei der Qualitätssicherung zu achten ist und die ihm zur Orientierung während des Untersuchungsprozesses dienen.

Gütekriterien oder Qualitätsstandards unterscheiden sich von Leitlinien durch die Definition nachprüfbarer Anforderungen für jede einzelne Leitlinie. Es leuchtet ein, dass es schwer fällt, für Leitlinien wie *Kontextangemessenheit der Methode* eindeutige Anforderungen zu formulieren, da wir es bei einer Beurteilung der Kontextangemessenheit mit "weichen Daten" zu tun haben. Eine Standardisierung hilft hier nicht weiter. Stattdessen muss als erster Schritt verlangt werden, dass bei einem Forschungsprojekt das Vorgehen so ausführlich und lückenlos dokumentiert wird, dass sich der Leser bzw. Nutzer ein eigenes Urteil über die Erfüllung der Qualitätsleitlinien bilden kann.

An Stelle der Qualitäts-Indizes, wie sie in der psychometrischen Diagnostik und beim Einsatz quantitativer Forschungsmethoden üblich sind, sollte die Qualitätsbeurteilung in der qualitativen Diagnostik und Forschung mit Hilfe einer *gutachterlichen Stellungnahme* erfolgen, in der kritisch abzuwägen ist, wieweit die Qualitätsleitlinien erfüllt wurden. Als zusätzliche Hilfe können hierbei *Checklisten* eingesetzt werden, in denen pro Leitlinie die erreichte Qualität durch *eine Urteilsskala* [nicht / teilweise / vollständig erfüllt / trifft nicht zu] erfasst wird.

Die dargestellten Leitlinien dienen dementsprechend zum einen als Anforderungskatalog für die Planung und Durchführung eines Projektes (*Qualitätssicherung*). Zum anderen haben sie die Funktion eines Pflichtenkatalogs für die Darstellung der Methoden und Ergebnisse im Forschungsbericht und ermöglichen damit dem Leser die Beurteilung der Seriosität einer Forschungsarbeit (*Qualitätskontrolle*).

5 Anregungen für die Diskussion

- ◆ Diskutieren Sie die Unterschiede und Gemeinsamkeiten der vier Wahrheitstheorien. Welche Wahrheitstheorie kommt Ihrer eigenen Vorstellung von Wahrheit im Bezug auf psychosoziale Phänomene am nächsten?
- ◆ Weshalb sind die Konstruktionen im Forschungsprozess und die Gegenstandsentscheidung im Interview wichtig für die Qualitätssicherung qualitativer (und quantitativer???) Methoden?
- ◆ Diskutieren Sie die Bedeutung der klassischen Gütekriterien und ihre Anwendung auf qualitative Methoden. Weshalb sind die Test-Gütekriterien hier nicht angemessen und wie lassen sie sich für qualitative Methoden modifizieren?
- ◆ Diskutieren Sie das Problem der externen Validität am Beispiel der Diagnostik von Angstzuständen (a) mit Hilfe eines Fragebogens zur manifesten Angst und (b) mit Hilfe eines offenen Interviews.
- ◆ Diskutieren Sie die unterschiedlichen Qualitätsaspekte qualitativer Methoden. Wie lassen sich diese Qualitätsaspekte zur Qualitätssicherung heranziehen?
- ◆ Diskutieren Sie den Unterschied zwischen Leitlinien und Standards bzw. Kriterien für die Qualität qualitativer Methoden.

6 Literatur

- Flick, U. (2004). Zur Qualität qualitativer Forschung – Diskurse und Ansätze. In: U. Kuckartz & H. Grunenberg (Hrsg.), *Qualitative Datenanalyse computergestützt*, 43-63. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Habermas, J. (1981). *Theorie des kommunikativen Handelns*. Bd. 1 und 2. Frankfurt: Suhrkamp.
- Huang, Kuan-Tsae, Yang W. Lee, Richard Y. Wang (1999). *Quality Information and Knowledge*. Upper Saddle River, NJ: Prentice Hall PTR.
- Legewie, H. (1987). Interpretation und Validierung biographischer Interviews. In: G. Jüttemann & H. Thomae (Hrsg.), *Biographie und Psychologie*, 138-150. Heidelberg: Springer.
- Steinke, I. (1999). *Kriterien qualitativer Forschung. Ansätze zur Bewertung qualitativ-empirischer Sozialforschung*. Weinheim: Juventa
- Strübing, J. (2004). *Grounded Theory – Zur sozialtheoretischen und epistemologischen Fundierung des Verfahrens der empirisch begründeten Theoriebildung*. Wiesbaden: VS Verlag für Sozialwissenschaften (S. 75-89).