

# 1 ARGUMENTE UND BEISPIELE FÜR DEN NUTZEN VON PRAXISFORSCHUNG

Joachim König

»Welchen Stellenwert hat Soziale Arbeit in unserer Gesellschaft? Warum ist welche Form der Sozialen Arbeit wie wichtig? Wie gut ist Soziale Arbeit eigentlich?«

Praxisforschung kann – so die These dieses Buches – auf drei Ebenen einen Beitrag zur Beantwortung dieser drängenden Fragen leisten:

- *Soziale Arbeit als gesellschaftliches Teilsystem* sieht sich von anderen Bereichen, vor allem der Politik und der Ökonomie herausgefordert. Es geht dabei um Verteilungsfragen und um Verteilungskämpfe, in denen anscheinend nur eines zählt: eine schlüssige Antwort auf die Frage nach dem Wert dieses Teilbereichs Soziale Arbeit für die gesamte Gesellschaft. Im Zeichen der Verknappung öffentlicher Haushalte gerät selbstverständlich auch Soziale Arbeit zunehmend unter Legitimationsdruck und in den Strudel der überall geführten Kürzungs- und Streichungsdebatten. Differenzierte Evaluationskonzepte können PraktikerInnen jedoch inzwischen in die Lage versetzen, sinnvolle Nachweise der Wirtschaftlichkeit Sozialer Arbeit zu führen.
- *Soziale Arbeit als wissenschaftliche Disziplin* benötigt für ihre Theoriebildung dringend Wissen darüber, welche unterschiedlichen Funktionen, Wirkungen und Nutzen Soziale Arbeit im Hinblick auf die riesige Palette unterschiedlicher Zielgruppen hat. Auch in ihrem Verhältnis zu den Nachbardisziplinen sind im Diskurs um die Wissenschaftlichkeit Sozialer Arbeit in theoretischer Hinsicht Bemühungen entstanden – auch durch verstärkte Praxisevaluationen –, zur Theoriebildung der »Sozialarbeitswissenschaft« beizutragen.
- *Soziale Arbeit als Profession* besinnt sich seit einigen Jahren auf einen zuvor wenig berücksichtigten Teilbereich ihres methodischen Handelns: Indem Praxisforschung als ein hilfreiches Instrumentarium in das Alltagsgeschäft der Sozialen Arbeit integriert wird, kann sie basisnahe und gleichzeitig präzise Erkenntnisse liefern, worin im Einzelfall oder auf eine bestimmte Gruppe von AdressatInnen bezogen der Wert des beruflichen Handelns liegt. Im Rahmen der Professionalisierungsdebatte hat sich daher in den letzten Jahren verstärkt die Einsicht durchgesetzt, dass praxis-, lebenswelt- und handlungsorientierte Forschung in den verschiedenen Arbeitsbereichen der Sozialen Arbeit gezielte Beiträge zur Steigerung ihrer Fachlichkeit und damit zur Entwicklung neuer Standards methodischen Handelns leisten kann (vgl. dazu zunächst Heiner et al. 1994).

Mindestens auf diesen drei Ebenen entsteht also – so die Behauptung dieses Buches – ein Nutzen für die Soziale Arbeit, wenn sie versucht, systematisches Wissen über sich selbst zu generieren. Wie aber konkretisiert sich dies im Alltag, in der Praxis der Sozialen Arbeit selbst?

*Zurück zu unserem Beispiel:*

Bereits die ersten Auswertungen ergeben teilweise erhebliche Unterschiede zwischen den Einschätzungen von Wirkung und Zufriedenheit auf beiden Seiten. Einrichtungsintern wird daraufhin zunächst an einer einfachen, für alle verständlichen Veröffentlichung dieser Ergebnisse in Form von übersichtlichen und grafisch ansprechend gestalteten Plakaten gearbeitet, die im Freizeitraum des Hauses ausgestellt werden. Drei Wochen später findet eine Hausversammlung statt, auf der die Einschätzungen unter den Bewohnern und den MitarbeiterInnen diskutiert, gemeinsam geklärt und in Bezug auf mögliche Veränderungen konkretisiert werden. Schnell ergeben sich zentrale Knackpunkte und Schlüsselprozesse, die – für alle nachvollziehbar – den Ansatzpunkt für konzeptionelle Veränderungen und Weiterentwicklungen bilden. Eine kleine »Reformkommission« unter Beteiligung von Bewohnern und MitarbeiterInnen erarbeitet in der Folge Vorschläge zur Verbesserung der Angebotsstruktur im Freizeit- und gruppenpädagogischen Bereich.

Auf einer zweiten Schiene wird ein detaillierter Bericht über Ergebnisse und Schlussfolgerungen aus diesem Praxisforschungsprojekt erarbeitet. Dieser Bericht soll dann Grundlage für die Formulierung einer Qualifizierungsoffensive gegenüber dem Kostenträger auf Landesebene sein. Hier geht es vor allem um strukturelle Weiterentwicklung der Angebote, die teilweise kostenneutral zu realisieren sind, teilweise jedoch mit einer notwendigen Erhöhung der Tagessätze verbunden sind. Es wird vorgeschlagen, eine zusätzliche Förderung dieser neuen Angebote im Rahmen eines Modellprojekts des Bundes zu beantragen. Auch die Einschätzungen und Vorschläge der hausinternen »Reformkommission« werden in den Bericht aufgenommen. Nach einem Vorgespräch mit dem zuständigen Referenten im Landesverband findet außerdem im Rahmen der alljährlichen Landesversammlung eine 30-minütige Präsentation des Verlaufs und der Ergebnisse dieses Verbesserungsprozesses statt.

Zentrale Botschaft an die VertreterInnen aus Verbänden und Sozialpolitik ist dabei die fachliche Legitimation und Profilierung der therapeutischen und pädagogischen Angebote und die damit verbundene Reformbereitschaft der Einrichtung, auch im Hinblick auf die drängenden Fragen nach der Wirtschaftlichkeit künftiger Angebote und Dienstleistungen im Bereich der stationären Suchthilfe.

Diese Herangehensweise kann als exemplarisch gelten für viele Ansätze der Praxisforschung. Und dieses Beispiel beinhaltet auch alle wesentlichen Nutzen, die Praxisforschung für das Alltagsgeschäft der Fachkräfte und Verantwortlichen in der Sozialen Arbeit haben kann. Systematisch betrachtet handelt es sich dabei immer wieder im Wesentlichen um fünf Perspektiven und Herausforderungen:

- **Kontrolle:** Mithilfe von Praxisforschung sind soziale Dienste in der Lage, sich einer differenzierten Erwartungs-Erfolgs-Kontrolle zu unterziehen. Eine leistungsbezogene Prüfung kann Bewertungsgrundlagen schaffen, um Erfolg und Misserfolg auf der fachlichen und auf der politischen Ebene diskutierbar zu machen, sowohl mit Blick auf die Effektivität von Maßnahmen, Projekten und Angeboten (also bezogen auf die Frage nach der Zielerreichung) als auch

auf deren Effizienz (also bezogen auf das Verhältnis zwischen Aufwand und Wirkung von Praxis).

- *Aufklärung*: Nicht nur der finanzielle, auch der Problemdruck steigt in den Feldern der Sozialen Arbeit. Häufig besteht großer Klärungsbedarf angesichts praktischer Probleme im Alltagsgeschäft oder auch darüber hinaus mit Blick auf konzeptionelle Fragen. Durch geeignete Ansätze der Praxisforschung ergeben sich für Einrichtungen und Fachkräfte Möglichkeiten, Beiträge zur Strukturierung, zu mehr Klarheit in der Unübersichtlichkeit und Komplexität alltäglicher Aufgabenstellungen zu leisten – etwa durch die Rekonstruktion von Interventionsverläufen oder von Ursache-Wirkungs-Zusammenhängen.
- *Qualifizierung*: Seit vielen Jahren wird im Rahmen der Professionalisierungsdebatte in der Sozialen Arbeit die Forderung nach dem »Ende der Beliebigkeit« thematisiert und in dem Zusammenhang die Weiterentwicklung des methodischen Handelns betrieben. Auch Kompetenzen im Bereich der Praxisforschung können hier dazu beitragen, die Fachlichkeit der Mitarbeitenden zu optimieren. Aus dieser Einsicht können zudem sinnvolle Beiträge zur Personalentwicklung oder für neue Weiterbildungskonzepte in sozialen Organisationen und Verbänden entstehen.
- *Innovation*: Wie seit langer Zeit im Bereich der Industrie und in Organisationen anderer Humandienstleistungen sind auch in vielen Einrichtungen der Sozialen Arbeit sogenannte »kontinuierliche Verbesserungsprozesse« inzwischen zu praktikablen und effektiven Instrumenten nicht nur der Qualitätssicherung und -entwicklung geworden. Auch im Hinblick auf die Verbesserung struktureller Bedingungen alltäglicher Handlungsabläufe kann Praxisforschung nützlich sein und innovativ wirken, zur Erneuerung von Strukturen und Hilfeprozessen beitragen, konzeptionelle Weiterentwicklung initiieren oder auch bei der Entwicklung von Alleinstellungsmerkmalen hilfreich sein.
- *Legitimierung*: Die Konkurrenz der Anbieter auf dem Sozialmarkt wächst. Gleichzeitig ist in vielen Bereichen eine massive Reduzierung der öffentlichen Finanzierung zu beobachten. Dies hat zwangsläufig eine Erhöhung des Legitimationsdrucks aus Sicht der Anbieter zur Folge. Auch hier ist Praxisforschung in der Lage, durch den Nachweis von Effekten und Wirkungen oder die Dokumentation der Effizienz von Maßnahmen belastbare Aussagen über die Qualität Sozialer Arbeit im Einzelnen zu generieren. So können nach außen, etwa Kostenträgern oder der politischen Öffentlichkeit gegenüber, tragfähige Aussagen zur »Daseinsberechtigung« Sozialer Arbeit in fachlicher Hinsicht, aber auch im betriebs- und volkswirtschaftlichen Sinne gemacht werden.

## 2 WISSENSCHAFTSTHEORETISCHE VORÜBERLEGUNGEN

*Karl-Hermann Rechberg*

### Was Sie in diesem Kapitel lernen können

Um solide Praxisforschung zu betreiben, genügt es nicht, in den wissenschaftlichen Methodenbaukasten zu greifen. Um sich wissenschaftlicher Methoden korrekt zu bedienen, ist es notwendig, sich aus wissenschaftstheoretischer Sicht darüber im Klaren zu sein, welche Art von Daten man generiert, wozu diese dienen können und wozu nicht. Hierzu soll das folgende Kapitel einen Einstieg bieten.

Die Forderung an empirische Sozialforschung lautet, ihre Aussagen auf der Basis von Datenerhebung zu entwickeln und überprüfbar zu machen. Ihre Ergebnisse sind dabei stets von verschiedenen Aspekten geprägt. Zwei Beispiele sollen dies illustrieren:

- *Die fachliche Kompetenz der Forschenden:* Führt der Forscher ein Interview, werden die Ergebnisse auch von seinem Geschick der Gesprächsführung beeinflusst, den Gesprächspartner zum Reden zu animieren, ohne ihm dabei bestimmte erwünschte Aussagen zu suggerieren.
- *Die Möglichkeiten und Grenzen der angewendeten Forschungsmethoden:* Mithilfe des Mittelwertes können die Altersstrukturen verschiedener Gemeinden nur begrenzt verglichen werden. Gleiche Mittelwerte zweier Gemeinden sagen beispielsweise noch nichts über die Streuung ihrer Altersstruktur aus. Eine Gemeinde mit dem Namen Breitenkirchen mit einer stark gestreuten Altersstruktur könnte aus vielen Senioren und vielen Kindern und Jugendlichen bestehen, eine Gemeinde mit Namen Schmalenkirchen dagegen ausschließlich aus Personen mittleren Alters. Beide könnten jedoch den gleichen Altersmittelwert aufweisen. Das Beispiel zeigt: Der Mittelwert ermöglicht einerseits einen Vergleich der beiden Gemeinden, ist jedoch kein gleichwertiger Ersatz für alle einzelnen Altersangaben, da ihm unter anderem die Information über die Streuung der Altersstruktur fehlt.

Im zweiten Beispiel wird deutlich, dass beispielsweise statistische Kennziffern in erster Linie dazu dienen, die Sicht auf die komplexe Realität auf ein überschaubares Maß zu reduzieren. Oft wird es erst auf diese Weise möglich, sich in einer komplexen Welt zu orientieren und beispielsweise politische Entscheidungen zu treffen oder das praktische Handeln weiterzuentwickeln.

Bei der Koppelung solcher politischer Entscheidungen an wissenschaftliche Befunde sollte jedoch unter anderem bedacht werden, dass diese Befunde nicht

notwendigerweise eindeutig sind und damit die Unsicherheiten der Entscheidung nicht völlig ausgeräumt sein müssen. Eine kritische Auseinandersetzung mit sogenannten evidenzbasierten politischen Entscheidungen ist bei Müller und Waldow (2011) nachzulesen.

Grundsätzlich ist bei der Weiterentwicklung der Praxis auf der Basis von empirischen Untersuchungen unter anderem zu beachten, wie gewährleistet werden kann, dass die wissenschaftlichen Ergebnisse überhaupt für die Praxis nutzbar sind. Verschiedene Strategien, diesen Transfer zu leisten, können am Beispiel Evaluation von Schulunterricht bei Gräsel (2010) nachgelesen werden.

Unabhängig von diesen praktischen Überlegungen ist darauf zu achten, dass die wissenschaftlichen Ergebnisse, die Entscheidung und Evaluation zugrunde liegen, solide sind. Dies hängt unter anderem davon ab, ob das wissenschaftliche Vorgehen adäquat ist. Hierzu stellt dieses Kapitel im Folgenden einige wissenschaftstheoretische Überlegungen an.

## 2.1 Systematische Gestaltung des Forschungsvorhabens

Die Gestaltung der Forschungslogik und insbesondere die Auswahl der Methoden sollten danach ausgerichtet sein, inwiefern sie zum Anliegen einer Untersuchung passen. Um hierfür eine Orientierung zu geben, wird in diesem Abschnitt die Systematisierung der zur Verfügung stehenden Möglichkeiten zunächst nach zwei Forschungsparadigmen gegliedert. Dem wird am Ende des Abschnitts eine weitere alternative Systematisierung gegenübergestellt.

In der empirischen Sozialforschung haben sich inzwischen zwei Traditionen herausgebildet: Die quantitative und die qualitative Sozialforschung werden häufig als zwei Paradigmen bezeichnet, da sie unterschiedlichen Forschungslogiken folgen. In Deutschland werden sie daher in der Literatur häufig getrennt behandelt. Dies ist jedoch nicht zwingend nötig, wie beispielsweise Werke aus der Fachliteratur zeigen, die beide Paradigmen einschließen und Gemeinsamkeiten sowie Ergänzungen aufzeigen (vgl. Baur und Blasius 2014; Bortz und Döring 2006).

Das Paradigma der quantitativen Sozialforschung stellt in den Vordergrund, Merkmale der Realität zu operationalisieren, also in zählbare Einheiten zu überführen. Die Zufriedenheit von Klienten der sozialen Arbeit wird beispielsweise mit Fragebogen-Items abgefragt, mithilfe derer die Klienten auf einer Skala angeben, wie sehr sie mit einer Leistung zufrieden waren. Hat ein Klient sich für ein »zufrieden« entschieden, wird diese Angabe beispielsweise mit der Zahl 1 codiert. Hat ein anderer Klient mit »nicht zufrieden« geantwortet, könnte dies in den Messwert 3 überführt werden. Die Zufriedenheit dieser Zweiergruppe kann mit einem Mittelwert von 2 ausgedrückt werden.

Aussagen zielen dabei auf mengenmäßige Repräsentativität. Das bedeutet: Wenn man aus einer Bevölkerungsgruppe eine Stichprobe zieht und diese quantitativ analysiert, werden bestimmte Strategien angewendet, welche die Wahrscheinlichkeit erhöhen, dass die Ergebnisse der Stichprobe mit denen der gesamten Bevölkerungsgruppe vergleichbar sind.

Forschungsvorhaben werden in dieser Tradition üblicherweise linear durchgeplant: Beispielsweise wird zuerst ein Fragebogen entworfen, der alle interessierenden Fragen enthält. Mit diesem wird anschließend eine Erhebung vorgenommen. Nach Aufbereitung der Daten werden diese schließlich ausgewertet. Hierbei ist wesentlich, dass die Ergebnisse vom Forscher unabhängig sind. Zu diesem Zweck bedient sich dieser Forschungszweig statistischer Methoden, die möglichst objektive Ergebnisse zum Ziel haben, mithilfe von Kennwerten Komplexität reduzieren und Vergleichbarkeit herstellen. Die Ergebnisse können schließlich in Diagrammen dargestellt werden.

Das Paradigma der qualitativen Sozialforschung arbeitet mit nicht-numerischem Material, wie vor allem Texten, aber auch Fotografien, Videos usw. Die Datenerhebung ist an vielen Details interessiert, die sich aus dem Datenmaterial selbst herausbilden sollen anstatt durch die Methode vorgegebene Merkmale zu messen. Beispielsweise werden Klienten Sozialer Arbeit in einem Interview offen gefragt, was sie an einer Leistung zufrieden gestellt oder gestört hat. Die Antworten können bestehende Erwartungen erfüllen, wenn Klienten beispielsweise sagen, dass sie Hilfe erhalten hätten. Es können aber auch überraschende Aspekte hinzukommen, beispielsweise dass es im Wartezimmer nicht so übel gerochen hat, wie es sonst oft der Fall sei.

Aussagen zielen dabei auf strukturelle Repräsentanz. Das bedeutet: In den Ergebnissen sollen die wesentlichen Aspekte enthalten sein, die es im Hinblick auf eine Fragestellung zu unterscheiden gilt. Es sollten beispielsweise die wesentlichen Zufriedenheitsaspekte der Klienten zusammengestellt werden. Damit wird keine Aussage darüber getroffen, wie häufig sie vorkommen.

Häufig werden Forschungsvorhaben in dieser Tradition nicht linear, sondern zirkulär durchgeführt. Das bedeutet, dass beispielsweise erst nach Erhebung und Auswertung einiger erster Fälle eine Informationsgrundlage besteht, auf deren Basis die weiteren Fälle ausgewählt werden (Theoretisches Sampling). Stellt man beispielsweise nach Erhebung der ersten fünf Fälle fest, dass schon lange bekannte Klienten ihre Zufriedenheit an völlig anderen Kriterien festmachen als solche, die eine Einrichtung zum ersten Mal aufsuchen, würde man diese beiden Extreme in den folgenden Erhebungen gezielter untersuchen. Ein Vergleich der Ergebnisse kann anhand bestimmter, im Auswertungsprozess herausgearbeiteter Kategorien vorgenommen werden. Diese entstehen jedoch grundsätzlich in einem interpretativen Prozess, wodurch die Subjektivität der Forschenden nicht gänzlich ausgeschaltet werden kann. Die qualitative Sozialforschung versucht diesen Umstand zwar einerseits durch verschiedene Verfahren der intersubjektiven Korrektur zu kompensieren. Beispielsweise interpretieren mehrere Forscher das gleiche Datenmaterial und vergleichen anschließend ihre Ergebnisse. Andererseits begreift sie Subjektivität auch als eine Ressource, die ein Verstehen der erforschten Subjekte erst möglich macht.

Neben einer paradigmatischen Unterscheidung in quantitative und qualitative Methoden der Datenerhebung besteht die Möglichkeit, Methoden in die Kategorien beschreibend, hypothesengenerierend oder hypothesenüberprüfend einzuteilen. Man könnte beispielsweise Klienten in einem Fragebogen folgende offene Frage stellen: »Was hat Ihnen in der Beratung nicht gefallen?«. Aus den Antwort-

ten lassen sich im qualitativen Sinne Kategorien bilden. Die Antworten »finstere Räume« und »alles war so dunkel« lassen sich beispielsweise der Kategorie »Raumhelligkeit« zuordnen. Diese Kategorien können in einem zweiten Schritt quantitativ auf die Frage hin ausgewertet werden, welche Kategorie besonders häufig genannt wurde. Wird die Kategorie »Raumhelligkeit« von 65 % der Befragten genannt, lässt sich auf dieser Grundlage die Hypothese generieren, dass diese Kategorie den Befragten besonders wichtig sein könnte. Hier wird also eine qualitative Erhebungsmethode quantitativ ausgewertet. Unabhängig von dieser Unterscheidung könnte man jedoch die beschriebene Forschungsstrategie als hypothesengenerierend bezeichnen. Zur näheren Erläuterung von beschreibenden und hypothesenüberprüfenden Strategien siehe Kapitel 4.8.

Unabhängig davon, nach welchem Paradigma ein Forscher arbeitet, und unabhängig davon, ob er beschreibend, hypothesengenerierend oder -überprüfend arbeiten möchte, kann es sich lohnen, unterschiedliche Methoden zu kombinieren, wie es bei der sog. Triangulation bzw. den Mixed-Methods-Designs gemacht wird (vgl. Flick 2011; Kuckartz 2014a): Werden innerhalb eines einzigen Forschungsvorhabens mehrere Datenerhebungen vorgenommen, so können dabei Verfahren aus verschiedenen Forschungstraditionen angewendet und aufeinander bezogen werden. Die bereits angeführten qualitativen bzw. hypothesengenerierenden Interviews mit Klienten der Sozialen Arbeit könnten beispielsweise dazu dienen, Aspekte der Zufriedenheit zu sammeln. In einem zweiten Schritt könnten diese Aspekte mithilfe eines standardisierten Fragebogens bei weiteren Klienten abgefragt werden, um anschließend im Sinne der quantitativen Forschung auszuzählen, welche Wünsche besonders häufig genannt werden bzw. zu beschreiben, wie die Wünsche mengenmäßig unter den Befragten verteilt sind.

## 2.2 Gütekriterien empirischer Sozialforschung

Die empirische Sozialforschung versucht, mit sogenannten Gütekriterien Standards zu finden, an denen die Qualität von Forschungsergebnissen abgelesen werden kann.

### 2.2.1 Gütekriterien quantitativer Forschung

Die quantitative Sozialforschung unterscheidet in Bezug auf die Qualität der Erstellung und Anwendung ihrer Messinstrumente die folgenden Gütekriterien (vgl. Krebs und Menold 2014; Bortz und Döring 2006):

- **Objektivität:** Eine Datenerhebung ist dann objektiv, wenn sie unabhängig vom Forscher, der sie durchführt, bei gleicher erhobener Personengruppe immer zum selben Ergebnis gelangt. Dies soll durch die Standardisierung von Fragebögen gewährleistet werden sowie dadurch, dass die Befragungssituation die Befragten möglichst nicht beeinflusst. Die Datenerhebung, -aufbereitung und -auswertung muss ausführlich dokumentiert werden, um ihre Objektivität nachvollziehbar zu machen, insbesondere wenn Störungen wie

fehlende Rückläufe auftreten, mit denen die Forscher irgendwie umgehen müssen. Die Interpretation der Ergebnisse kann nicht vollständig objektiv sein, da sie den Werturteilen der interpretierenden Person unterliegt.

- **Reliabilität:** Sie gibt die Messgenauigkeit des Erhebungsinstruments an. Sie zeigt sich beispielsweise darin, dass Messergebnisse möglichst exakt auch bei weiteren Messungen wieder eintreten, also reproduziert werden können. Man geht davon aus, dass ein Messwert aus einem wahren Wert und einem Messfehler besteht. Solche Messfehler können dadurch auftreten, dass die Befragten beim Ausfüllen eines Fragebogens durch Müdigkeit oder Störungen beeinträchtigt sind oder raten, wenn sie eine Frage nicht eindeutig beantworten können, oder sogar bewusst lügen. Es ist also grundsätzlich vom Messfehler eines Instruments auszugehen. Je kleiner der Messfehler ist, als desto reliabler gelten die Ergebnisse. Es gibt mehrere Reliabilitätstestverfahren, welche die Messergebnisse mehrerer Messungen vergleichen, um von deren Schwankungen auf die Größe des Messfehlers zu schließen. Dabei sind Tests, welche die Reliabilität einzelner Fragebogenitems messen, von solchen zu unterscheiden, welche die Reliabilität ganzer Item-Sammlungen messen. Zu den einzelnen Messverfahren sei beispielsweise auf die ausführliche Darstellung von Testgütekriterien bei Bortz und Döring (2006) verwiesen.
- **Validität:** Sie gibt an, in welchem Maß ein Erhebungsinstrument genau das misst, was es auch messen soll. Versucht ein Fragebogen mit der Frage: »Wie fühlen Sie sich jetzt nach dem Beratungsgespräch?«, die Zufriedenheit eines Klienten mit einem sozialen Beratungsangebot zu messen, so könnte es sein, dass die emotionalen Persönlichkeitsmerkmale der Person die Validität des Ergebnisses verzerren: Eine grundsätzlich missmutige Person würde vermutlich tendenziell negativere Antworten geben als eine grundsätzlich fröhliche Person. Um die Validität zu bestimmen, werden Messdaten zu anderen Daten in Beziehung gesetzt, die etwas Ähnliches messen. Auch zu den Messverfahren der Validität ist die ausführliche Darstellung von Testgütekriterien bei Bortz und Döring (2006) sehr hilfreich.

Die Validität einer Messung hängt von der Reliabilität des Messinstruments ab, aber nicht umgekehrt: Ist eine Waage schlecht geeicht, ist ihre Reliabilität beeinträchtigt. Dann zeigt sie auch immer das falsche Messergebnis an. Das heißt: das Ergebnis ist nicht valide. Jedoch kann auch eine gut geeichte und damit reliable Waage ein fehlerhaftes Messergebnis anzeigen, wenn beim Messen jemand unbemerkt ein zusätzliches Gewicht mit auf die Waage stellt.

Neben den beschriebenen klassischen Gütekriterien geben die folgenden Aspekte Aufschluss über die Aussagekraft von quantitativen Untersuchungen (vgl. Krebs und Menold 2014; Bortz und Döring 2006):

- **Interne Validität:** Häufig gehört es zu den Ergebnissen von quantitativen Untersuchungen, dass Zusammenhänge zwischen mehreren Variablen berechnet werden. Beispielsweise wird ein statistischer Zusammenhang zwischen einem sozialen Kompetenztraining und einer Veränderung des sozialen Verhaltens seiner Absolventen festgestellt. Die interne Validität ist umso höher, je eher ausgeschlossen werden kann, dass dieser Zusammenhang auf andere Fakto-



ren als das Kompetenztraining zurückgeführt werden kann. Inwiefern dies gewährleistet ist, hängt vom Design der jeweiligen Untersuchung ab. Sogenannte experimentelle Designs versuchen durch einen stark geplanten Versuchsaufbau eine hohe interne Validität herzustellen. Dies geschieht jedoch oft auf Kosten der externen Validität.

- *Externe Validität:* Dieses Kriterium beschreibt, wie gut sich die Ergebnisse einer Untersuchung verallgemeinern lassen. Dies ist umso schlechter möglich, je mehr das Untersuchungsdesign von den üblichen Bedingungen des Alltags abweicht. Ein stark geplanter und standardisierter Versuchsaufbau, der wie bei einem Laborexperiment vorgibt, wie eine Situation ablaufen soll, steht in einem gewissen Widerspruch zum Alltag, bei dem sich Situationen zwar ähneln, aber doch in vielen Details unterscheiden. Die externe Validität sinkt ebenfalls, wenn die realisierte Stichprobe in wesentlichen Bestandteilen von der Grundgesamtheit abweicht, also mengenmäßig nicht repräsentativ ist. Zur mengenmäßigen Repräsentativität von Stichproben ist die gut verständliche Einführung von Mayer (2013) zu empfehlen. Das gilt nicht nur für Untersuchungen, die Zusammenhänge berechnen, sondern auch für die sogenannte beschreibende Statistik.
- *Statistische Signifikanz:* Untersucht man den Zusammenhang zwischen einem sozialen Kompetenztraining und dem sich verändernden sozialen Verhalten von zwei Teilnehmern im Vergleich zu zwei Nicht-Teilnehmern, so kann es sein, dass sich ein hoher Zusammenhang ergibt. Jedoch ist bei insgesamt vier Versuchspersonen die Wahrscheinlichkeit äußerst hoch, dass es sich um ein zufälliges Ergebnis handelt. Die Signifikanz ist ein statistisches Maß, das diese Wahrscheinlichkeit ausdrückt. Hoch signifikante Ergebnisse sind mit einer hohen Wahrscheinlichkeit nicht zufällig. Die Signifikanz ist insbesondere abhängig von der Menge der beteiligten Personen. Sie sollte bei der Interpretation von Zusammenhängen stets genannt werden. Zur ihrer Berechnung sei erneut auf die Darstellung bei Bortz und Döring (2006) hingewiesen.
- *Effektstärke:* Besteht wirklich ein hoch signifikanter Zusammenhang zwischen einem sozialen Kompetenztraining und der Verbesserung sozialen Verhaltens seiner Teilnehmer, so muss dies nicht bedeuten, dass diese Verbesserung eine hohe Relevanz besitzt. Die sogenannte Effektstärke beschreibt, in welcher Intensität sich ein Zusammenhang auswirkt. Wenn man die Veränderung des sozialen Verhaltens auf einer 10-Punkte-Skala misst, dann wäre eine Verbesserung um 0,3 Punkte ein verhältnismäßig niedriger Effekt. Dies kann beispielsweise darauf hindeuten, dass das Kompetenztraining zwar wirksam ist, aber für eine weitere Finanzierung nicht wirksam genug. Die Effektstärke sollte bei der Interpretation von Zusammenhängen ebenfalls genannt werden. Auch ihre Berechnung ist bei Bortz und Döring (2006) nachzulesen.

### 2.2.2 Gütekriterien qualitativer Forschung

Flick (2014; 2012) stellt die Diskussion um die Gütekriterien in der qualitativen Sozialforschung ausführlich dar. Sie hat bislang nicht zu einem Konsens geführt,

der mit den Standards der quantitativen Sozialforschung vergleichbar wäre. Jedoch lassen sich einige Zwischenergebnisse umreißen:

- Eine unmittelbare Übertragung der Gütekriterien Objektivität, Reliabilität und Validität aus dem quantitativen Paradigma erscheint aufgrund der veränderten Forschungsbedingungen nicht empfehlenswert. Versteht man beispielsweise Reliabilität als Reproduzierbarkeit von Ergebnissen und will man die Reliabilität eines narrativen Interviews damit belegen, dass dieselbe Person bei einem zweiten oder dritten Interview ihre Erzählung wiederholt, ist davon auszugehen, dass sich die Person ihre Erzählung »zurechtgelegt« hat und dadurch die Gefahr einer konstruierten Erzählung steigt.
- Eine häufig angewendete Form der Validierung qualitativer Ergebnisse besteht darin, dass sie den Beforschten vorgelegt und diese um Rückmeldung gebeten werden, ob sie den Ergebnissen zustimmen. Die Vorschläge unterscheiden sich jedoch darin, ob man beispielsweise jeder Person die Ergebnisse zu ihrem individuellen Interview vorlegt oder das Gesamtergebnis der Untersuchung. Ein viel größeres noch ungelöstes Problem ist jedoch, wie viel Zustimmung erforderlich ist, damit ein Ergebnis als valide gelten kann, sowie die Frage, wann Differenzen, die sich aus den unterschiedlichen Perspektiven von Beforschten und Forschern ergeben, nicht der Validität zu Lasten gelegt werden dürfen.
- Eine weitere Form der Validierung besteht darin, die Forschungsergebnisse Experten vorzulegen, um deren Zustimmung zu erfragen. Diese können entweder aus dem Forschungsfeld stammen oder von diesem völlig unabhängig sein. Aber auch hier stellt sich die Frage, wie viele und welche Differenzen Validität schmälern.
- Rekonstruktive Forschungsansätze wie beispielsweise die dokumentarische Methode nach Bohnsack (2003) basieren auf einem eigenen theoretischen Überbau. Eine Validierung besteht dann unter anderem darin, den Forschungsprozess und seine Ergebnisse in Beziehung zu diesem theoretischen Rahmen zu setzen. Auswertungsmethoden, die einen solchen Rahmen nicht besitzen, sind von diesem Ansatz jedoch ausgeschlossen. Standards dieses Vorgehens sind bei Bohnsack (2005) nachzulesen.
- Es wurde bereits versucht, eine Validierung von qualitativen Forschungsergebnissen durch Triangulation durchzuführen, das heißt, durch den Einsatz weiterer Methoden, zusätzlicher Daten, verschiedener Forscher im Erhebungsprozess oder Anwendung verschiedener Theorien. Die Ergebnisse einer qualitativen Methode sollten dabei durch die Ergebnisse beispielsweise einer quantitativen Methode bestätigt oder infrage gestellt werden. Dies ist jedoch dann problematisch, wenn Unterschiede in den Ergebnissen auch durch die unterschiedlichen Perspektiven ausgelöst werden können, wenn beispielsweise die qualitative Methode hypothesengenerierend und die quantitative hypothesenprüfend eingesetzt wurde.
- Als Konsens kann immerhin formuliert werden, dass eine ausführliche Forschungsdokumentation anzufertigen ist, die es Dritten nicht nur durch ihre Genauigkeit und Ausführlichkeit, sondern auch durch ihre Verständlichkeit