

## Die Gruppendiskussion als Evaluationsmethode – Entwicklungsgeschichte, Potenziale und Formen

*Susanne Mäder*<sup>1</sup>

**Zusammenfassung:** Der Artikel beabsichtigt, den derzeitigen Diskussionsstand zu Theorie und Praxis der Gruppendiskussion bezogen auf das Feld der Evaluation zusammenzufassen, die Stärken und Grenzen von Gruppendiskussionen in evaluativen Vorhaben zu skizzieren und eine weitere Fundierung der Gruppendiskussion in der Evaluation anzuregen. Nach einer definitorischen Klärung verwandter Begrifflichkeiten wie Gruppendiskussion, Focus Group und Gruppeninterview wird ein Überblick über die Entwicklung und Rezeption der Methode in der Literatur der Sozialforschung und Evaluation gegeben. Es wird herausgearbeitet, was Evaluation als Disziplin – im Vergleich zur Sozialforschung allgemein – ausmacht und welche besonderen Leistungspotenziale sich hieraus für die Gruppendiskussion in der Evaluation ergeben. Zwei idealtypische Formen der Gruppendiskussion werden kontrastiert und zentrale Gestaltungsparameter bei der Planung, Durchführung und Auswertung vorgestellt und diskutiert.

**Schlagwörter:** Gruppendiskussion, Gruppeninterview, Fokusgruppe, Qualitative Evaluationsmethoden

## Group Discussion as an Evaluation Method – Development, Potentials and Models

**Abstract:** This article aims to summarize the progress in the current discussion on theory and practice of group discussions in evaluation and thereby to professionalize the application of this technique in evaluation. After having clarified the definitions of related terms such as group discussion, focus group and group interview, an overview of the development and reception of group discussions in social science and evaluation is provided. It is shown what differentiates evaluation as a discipline when compared to social science in general. From these differentiating characteristics of evaluation result particular potentials for data collection and the design of the evaluation process applying group discussions. Finally, two ideal-typical forms of group discussions are contrasted with each other and vital elements for planning, implementation and analysis are discussed.

**Keywords:** Group Discussion, Group Interview, Focus Group, Qualitative Evaluation Methods

---

1 Univation, Institut für Evaluation, Köln

## 1. Einleitung

Das Interesse an der Methode der Gruppendiskussion in der Evaluation wächst. In der Evaluationspraxis werden Daten immer häufiger in gruppenförmigen Settings erhoben. So kommt die Methode mittlerweile in Evaluationsvorhaben verschiedener Anwendungs- und Politikfelder zum Einsatz, wie im Bereich der Gesundheitserziehung und -förderung, der Evaluation von Reformprozessen im Bildungs- und Sozialsystem, der Bewertung von Programmen in Schule und Schulsozialarbeit sowie der Hochschulevaluation. Evaluationen zeichnen sich durch eine ausgeprägte Praxis- und Nutzenorientierung, einen Bewertungsanspruch sowie eine Dialogorientierung aus. Diese charakteristischen Merkmale erfordern von Evaluierenden in hohem Maße kommunikative und interaktive Leistungen. Gruppenmethoden wie die Gruppendiskussion sind besonders geeignet, zur Erfüllung dieser Aufgaben einen Beitrag zu leisten. Gruppendiskussionen bieten die Möglichkeit, Sichtweisen, Einstellungen, Geschichten und Selbstrepräsentationen in einem gegebenen kulturellen und kollektiven Kontext zu erheben. Sie können Aufschluss darüber geben, wie Meinungen und Sichtweisen konstruiert sind und artikuliert werden.

Die methodologische und methodische Auseinandersetzung mit der Gruppendiskussion als Evaluationsmethode steht ungeachtet ihrer zunehmenden Beliebtheit jedoch noch am Anfang. Der Artikel beabsichtigt, den derzeitigen Diskussionsstand zu Theorie und Praxis der Gruppendiskussion bezogen auf das Feld der Evaluation zusammenzufassen, die Stärken und Grenzen von Gruppendiskussionen in evaluativen Vorhaben zu skizzieren und die professionelle Nutzung als Evaluationsmethode zu fördern.

Nach einer definitorischen Klärung verwandter Begrifflichkeiten wie *Gruppendiskussion*, *Focus Group* und *Gruppeninterview* (Kapitel 2) wird ein Überblick über die Entwicklung und Rezeption der Methode in der Literatur der Sozialforschung (Kapitel 3) und Evaluation (Kapitel 4) gegeben. In einem nächsten Schritt wird herausgearbeitet, was die Evaluation als Disziplin – im Vergleich zur Sozialforschung allgemein – ausmacht (Kapitel 5). Es wird gezeigt, dass sich insbesondere in der Evaluation für die Methode der Gruppendiskussion – als Setting zur Erhebung von Daten sowie als Instrument zur Gestaltung des Evaluationsprozesses – besondere Leistungspotenziale ergeben (Kapitel 6). Abschließend werden zwei idealtypische Formen der Gruppendiskussion kontrastiert und zentrale Gestaltungsparameter bei der Planung, Durchführung und Auswertung vorgestellt, über die in der Evaluationspraxis entschieden werden muss (Kapitel 7). Im Fazit (Kapitel 8) werden die Argumentationslinien zusammengefasst und die Notwendigkeit weiterer Forschungsbedarfe aufgezeigt.

## 2. Definitiorische und begriffliche Klärungen

Zunächst soll geklärt werden, was unter einer Gruppendiskussion zu verstehen ist und wie benachbarte in der Literatur verwendete Begriffe einzuordnen sind. Ausgehend von unterschiedlichen Entstehungskontexten und länderspezifischen Entwicklungen finden sich in der Literatur mehrere Bezeichnungen für die Methode. Im deutschsprachigen Raum ist die Benennung *Gruppendiskussion* am weitesten verbreitet; stärker strukturierte Varianten, die sich mehr an der angelsächsischen Tradition einer fokussierten, die Gruppen steuernden Moderation orientieren, werden teilweise auch als Gruppeninterviews bezeichnet (vgl. Flick 2004). Überwiegend wird unter einem Gruppeninterview jedoch die gleichzeitige Befragung von mehreren Personen verstanden, wobei Interaktionen zwischen den Teilnehmenden im Gegensatz zur Gruppendiskussion nicht intendiert sind (vgl. Kromrey 1994; Lamnek 2005; Loos/Schäffer 2001). In der angelsächsischen Literatur trifft man auf die Bezeichnungen *group discussion*, *group interview* und *focus group*, wobei sich in letzter Zeit der Begriff *focus group* durchgesetzt zu haben scheint. Im angelsächsischen Verständnis umfassen *focus groups* sowohl strukturierte als auch eher selbstläufige Varianten von Gruppendiskussionen, während die Bezeichnung im deutschsprachigen Raum für strukturierte Formen des Verfahrens steht. Gruppendiskussionen werden in diesem Artikel in Anlehnung an Morgan (1997) als Erhebungsmethode definiert, die Daten durch die Interaktion der Gruppenmitglieder gewinnt, wobei die Thematik durch das Interesse der Forschenden bestimmt ist. Diese Definition erlaubt eine Abgrenzung zu Gruppenmethoden, die keine Interaktion zwischen den Teilnehmenden vorsehen (wie Delphi-Technik und Gruppeninterview), und zu solchen, die nicht durch die Forschenden gesteuert werden, wie natürliche Gruppengespräche (vgl. Lamnek 2005). Zudem ist die gewählte Definition so weit gefasst, dass verschiedene Varianten von Gruppendiskussionen eingeschlossen werden können (z.B. strukturierte bis selbstläufige Formate).

Im Folgenden wird zunächst ein kurzer Überblick über die Entwicklung und die Formen der Gruppendiskussion in der Sozialforschung gegeben. Das anschließende Kapitel geht speziell auf die Diskussionslinien, Einsatzfelder und Anwendungsvarianten der Methode für das Feld der Evaluation ein.

## 3. Entwicklung und Formen der Gruppendiskussion in der Sozialforschung

Im angelsächsischen Raum wurden Gruppendiskussionen in den 1940er Jahren zunächst von Robert King Merton als Erhebungsinstrument zur Erforschung der Wirkungen von Propagandamaßnahmen der US-Regierung während des 2. Weltkriegs eingesetzt (vgl. Merton 1987; Merton/Lowenthal/Kendall 1956, 1994; Merton/Kendall 1979). Mittels fokussierter (Gruppen-)Interviews wurde versucht, Reaktionsmuster und Einstellungstypen von Zuschauern zu identifizieren. Die Gruppen-

diskussionen fungierten hier als ergänzende Erhebungen im Rahmen großangelegter Surveyuntersuchungen. Trotz ihres sozialwissenschaftlichen Ursprungs wurde die Gruppendiskussion im angelsächsischen Raum zwischen 1950 und 1970 vorrangig als Instrument in der Marktforschung eingesetzt. Seit den 1980er Jahren wird sie insbesondere im Bereich der Gesundheitserziehung und Gesundheitsprävention innerhalb der Aktionsforschung genutzt sowie in der Gesundheitsforschung im Allgemeinen. Ausgelöst durch die Vorreiter Richard A. Krueger und David Morgan (vgl. Krueger 1994; Morgan 1988) hat die Gruppendiskussion in den Sozialwissenschaften in den 1990er Jahren enorm an Popularität gewonnen (vgl. Bloor/Frankland/Thomas/Robson 2001; Stewart/Shamdasani 1990). Dies führte zu einer substantiellen Anzahl an Publikationen in verschiedenen Disziplinen wie Gesundheitsforschung, Erziehungswissenschaft, Kommunikations- und Medienforschung, Feminismusforschung, Soziologie, Bildungsforschung, Behindertenforschung und Sozialarbeitsforschung (vgl. Bay-Cheng/Livingston/Fava 2011; Kroll/Barbour/Harris 2007; Liamputtong 2011; Linhorst 2002; Peek/Fothergill 2009; Schilder/Tomov/Mladenova/Mayeya/Jenkins/Gulbinat/Manderscheid/Baingana/Whiteford/Khandelval/Minoletti/Mubbashar/Srinivasa/Parameshvara/Baba/Townsend/Sakuta 2004; Willis/Green/Daly/Williamson/Bandyopadhyay 2009).

Als theoretische Bezugspunkte für die Methodologie der Gruppendiskussionen (vgl. Liamputtong 2011) werden in der englischsprachigen Literatur der symbolische Interaktionismus (vgl. Morgan 2012), der Feminismus (vgl. Madriz 2003; Robinson 2009) und die kritische pädagogische Praxis nach Paolo Freire und Jonathon Kozol angeführt (vgl. Kamberelis/Dimitriadis 2008, 2011). Die angelsächsische Diskussion über die Methode ist im Vergleich zur Entwicklung im deutschsprachigen Raum jedoch weniger von einzelnen akademischen Schulen und durch diese begründete Modelle der Gruppendiskussion geprägt. Anzutreffen ist vielmehr eine Offenheit für unterschiedliche methodische und konzeptionelle Vorgehensweisen. So werden beispielsweise Vor- und Nachteile bezüglich der Zusammensetzung von bestehenden Gruppen und künstlichen Gruppen erörtert und in Abhängigkeit von den jeweiligen Forschungsintentionen unterschiedliche Varianten propagiert. Entgegen des zu der Zeit vorherrschenden Trends, künstliche Gruppen zusammenzustellen, entscheidet sich Kitzinger (1994) zum Beispiel für bestehende Gruppen, um eine möglichst ‚natürliche‘ und authentische Diskussion der Themen zu erreichen. Da sich die Beiträge der Teilnehmenden auf gemeinsame Ereignisse eines geteilten Lebens- und Erfahrungszusammenhangs beziehen, könnten die geäußerten Sichtweisen und präsentierten Erfahrungen der anderen Teilnehmenden kritisch hinterfragt und ergänzt werden. Außerdem stellten die bestehenden Gruppen einen sozialen Kontext dar, in dem sich die Einstellungen und Ideen der interessierenden Personen formieren und entwickeln. Folgende zentrale Merkmale und Stärken der Gruppendiskussion, die allesamt aus dem interaktiven Charakter der Methode erwachsen, werden herausgestellt (vgl. Wilkinson 1998a, b; Morgan 2012): Gruppendiskussionen ermöglichen einen Zugang zu der Sprache, den Anliegen und Konzepten der Teilnehmenden. Sie böten eine Gelegenheit „to observe the process of sensemaking“ (vgl. Wilkinson 1998a: 193), indem untersucht werden könne, wie Meinungen artikuliert, modifiziert, verhandelt und verteidigt werden (vgl. Kitzinger 1994). Eine

Besonderheit von Gruppendiskussionen wird in der Möglichkeit gesehen, durch eine Analyse der interaktiven Prozesse die Co-Konstruktion von Bedeutung zu explorieren („co-construction of meaning“, ein Terminus, der erstmals durch Wilkinson 1998b eingeführt wurde). Nach Morgan (2012), der die Prozesse der Bedeutungsgenerierung anhand von Transkripten analysiert, wird Bedeutung in einem ersten Schritt durch „sharing and comparing“ (Morgan 2012: 164) hergestellt. In dieser Phase erweitern die Teilnehmenden vorangegangene Statements um ähnliche bzw. unterschiedliche Erfahrungen und elaborieren so das Thema (z.B. individuelle Erfahrungen mit Herzinfarkten). Im Anschluss überführen die Teilnehmenden die erarbeiteten inhaltlichen Elemente in abstraktere Kategorien und Konzepte („Organizing and Conceptualizing“, Morgan 2012: 170, z.B. Gründe und Risikofaktoren im Zusammenhang mit Herzinfarkten).

In Deutschland ist die Entwicklungsgeschichte der Methode – weitgehend ohne Bezug zur angelsächsischen Literatur – eng verankert mit sozialwissenschaftlichen Traditionen und ihren unterschiedlichen theoretischen Ansätzen. Im Wesentlichen lassen sich vier als Modelle klassifizierte Ansätze identifizieren:

- Die Gruppendiskussion zur Ermittlung individueller nicht öffentlicher Meinungen: Diese Variante der Gruppendiskussion wurde im Rahmen eines großangelegten ‚Gruppeninstruments‘ in den Jahren 1950-51 zur Erfassung des politischen Bewusstseins der Deutschen durch Friedrich Pollock, Mitbegründer des Frankfurter Instituts für Sozialforschung, entwickelt (vgl. Pollock 1955). In kritischer Distanz zur Umfrageforschung wird in der Gruppendiskussion eine Methode gesehen, die den Bedingungen, unter welchen sich Meinungen in der Realität formen und verändern, besonders nahe kommt. So nimmt Pollock an, dass sich Meinungen zu Themen allgemeinen und öffentlichen Interesses nicht isoliert entwickeln, sondern „in ständiger Wechselbeziehung zwischen dem Einzelnen und der unmittelbar auf ihn einwirkenden Gesellschaft“ (Pollock 1955: 32). Durch die Gruppensituation sollte es gelingen, unter der Oberfläche die teils tabuisierten und latenten Meinungen der Deutschen zu identifizieren. Es wurde angenommen, dass die individuellen Meinungen mit der öffentlichen Meinung nicht deckungsgleich sind. Die Anlage der Gruppendiskussionen trug experimentelle Züge. So wurde beispielsweise als einheitlicher Grundreiz ein auf „psychologische Nervenpunkte“ hin konstruierter Brief eines amerikanischen Soldaten verlesen. Außerdem wurden vorab entwickelte sogenannte standardisierte Gegenargumente von der Moderation in die Diskussion eingebracht.
- Die Gruppendiskussion zur Ermittlung informeller Gruppenmeinungen: Aus der Perspektive von Werner Mangold (1973 und 1960), ebenfalls Mitarbeiter am Frankfurter Institut für Sozialforschung, lassen sich mit Gruppendiskussionen gerade nicht individuelle Meinungen, sondern ausschließlich Gruppenmeinungen zuverlässig erheben. Er macht am Beispiel von industriesoziologischen Studien deutlich, dass die Meinungsäußerung des Einzelnen durch die Gruppe beeinflusst wird, indem Einzelne ihre Meinung unterdrücken oder sich aus Angst, von der Gruppenmehrheit abzuweichen, Mehrheitsmeinungen anschließen. Die artikulierten Aussagen der Gruppenmitglieder werden daher als das

Produkt kollektiver Interaktionen und somit als Gruppenmeinung gesehen. Die Gruppenmeinung hat sich nach Mangold „in der Realität unter den Mitgliedern des betreffenden Kollektivs bereits ausgebildet“ (1973: 240) und wird in der Gruppendiskussion lediglich aktualisiert.

- Die Gruppendiskussion zur Ermittlung situationsabhängiger Gruppenmeinungen: Nach Nießen (1977), der die Position des symbolischen Interaktionismus vertritt, basiert jede Handlung auf einem eigenständigen situationspezifischen Interpretationsprozess. Die Bedeutungszuschreibungen der Subjekte sind an situationspezifische Aushandlungsprozesse gebunden, weshalb die Reichweite der in der Gruppendiskussion getätigten Aussagen nicht über die Gruppensituation hinausgehen kann. Um einen Transfer auf die realen Situationen zu erlauben, fordert Nießen, dass die an der Diskussion beteiligten Personen auch außerhalb der Untersuchungssituation eine Gruppe bilden (Realgruppen).
- Die Gruppendiskussion zur Ermittlung kollektiver Orientierungsmuster: Bohnsack (2004) kritisiert am Modell von Nießen, dass bei aller Prozesshaftigkeit die Strukturen aus dem Blick gerieten. Mit Bezug auf die Ergebnisse der Mediennutzungs- und Rezeptionsanalysen von David Morley betont er, dass die Gruppen im Diskurs „interpretative Codes“ (Morley 1996: 112) reproduzieren. Diese bezeichnen kollektive Sinn- und Bedeutungszuschreibungen, die zwischen Repräsentanten von Klassen bzw. Milieus existieren und sich dann aktualisieren, wenn Mitglieder der Gruppe zusammentreffen. Hiermit scheint ein Ausweg aus der methodologischen Sackgasse gegeben: Wenn es gelingt, derartige Codes, Muster oder Strukturen in Diskursen von Gruppen nachzuweisen, dann ließe sich die Zuverlässigkeit der Methode begründen. Genau diesen Weg beschreitet Bohnsack, indem er mittels der dokumentarischen Methode hinter dem „wörtlichen Sinngehalt“ einen „dokumentarischen Sinngehalt“ sichtbar machen will. Auf diese Weise soll deutlich werden, dass die Aussagen der Gruppenmitglieder durch ein „kollektives Orientierungsmuster“ miteinander verbunden sind, welches sich in verschiedenen Gruppen, die aus Angehörigen derselben Klassen oder Milieus bestehen, gleichartig entwickelt (vgl. Bohnsack/Nentwig-Gesemann 2010a und 2010b).

Bei zunehmender Popularität besteht bezüglich der methodologischen und methodischen Auseinandersetzung mit der Gruppendiskussion im Vergleich zu anderen Erhebungsmethoden wie dem Einzelinterview Nachholbedarf. Ihre weite Verbreitung und Beliebtheit in der Forschungspraxis selbst steht in einem Missverhältnis zu ihrer Rezeption und Berücksichtigung in der wissenschaftlich-methodischen Forschungsliteratur, in der sie eine eher randständige Position einnimmt. So konstatiert Lamnek (2005) in der Einleitung seiner Monographie zur Theorie und Praxis der Gruppendiskussion, dem einzigen klassischen deutschsprachigen Lehrbuch, dass die Methode in der Forschungspraxis „weitaus häufiger zur Anwendung kommt, als man dies aufgrund ihres inferioren Status in der methodologischen und methodischen Literatur annehmen würde“ (Lamnek 2005: 11).

In der quantitativ orientierten empirischen Forschungsliteratur wurden Gruppendiskussionsverfahren bislang eher marginal behandelt, wobei die skizzierten Nachteile des Verfahrens häufig die Vorteile überwiegen (vgl. Bortz/Döring 2002;

Friedrichs 1990). In der qualitativen Forschungsliteratur wird der Methode zwar mehr Beachtung geschenkt (vgl. Flick 2004; Lamnek 1995), im Vergleich zu anderen qualitativen Methoden war die Akzeptanz bislang jedoch eher gering. So äußert Morgan (2000) in einem Artikel zur Entwicklungsgeschichte der Gruppendiskussion: „By and large, however, established researchers with expertise in qualitative methodology have simply ignored focus groups“ (Morgan 2000: 145). Mittlerweile haben die Forschungen mit Gruppendiskussionen jedoch quantitativ zugenommen, wie Morgan in demselben Artikel weiterhin konstatiert. Auch an der Zunahme von Beiträgen über die Gruppendiskussion in Handbüchern ab dem Jahr 2000 ist ein verstärktes Interesse für die Methode zu bemerken (vgl. Bailey 2008; Colucci 2008; Flick/von Kardoff/Steinke 2005; Gaiser 2008; Kamberelis/Dimitriadis 2008, 2011; Kitzinger 2005; Macnaghten/Myers 2004; Morgan 2012; Smithson 2008).

In der Sozialforschung finden sich Studien, die die Gruppendiskussion als einzige Erhebungsmethode nutzen (vgl. Kitzinger 1993; McClelland/Newell 2008). Diese Forschungsprojekte intendieren meist, Fragestellungen aus der Perspektive zentraler Beteiligter zu untersuchen oder neue Forschungsgebiete, über die wenig Wissen vorliegt, zu explorieren. Außerdem werden Gruppendiskussionen als „supplementary source of data“ (Liamputtong 2011: 90) im Zusammenhang mit quantitativen Erhebungen eingesetzt. So erfolgt beispielsweise erst die explorative Erforschung eines Themas durch Gruppendiskussionen, um Informationen für die Planung weiterer Erhebungen zu erhalten (vgl. O'Brien 1993, Kitzinger 1994), oder es werden nachgängig zu anderen Erhebungen Gruppendiskussionen durchgeführt, mit dem Ziel, auffällige, unerwartete oder erklärungsbedürftige Ergebnisse vorangegangener Erhebungen zu erklären oder besser zu verstehen (vgl. Hennink 2007; Kitzinger 2005). Ebenso fungieren Gruppendiskussionen als gleichwertige Erhebungen in multimethodischen Untersuchungen. Beschrieben sind Kombinationsmöglichkeiten mit quantitativen Verfahren wie Fragebogenerhebungen (vgl. Makhoul/Nakkash 2007) oder auch qualitativen Methoden wie Beobachtungen und Interviews (vgl. Lambert/Loiselle 2008; Liamputtong 2011). Zum Beispiel kann in Gruppendiskussionen untersucht werden, wie in Einzelinterviews gewonnene interessante Einschätzungen und Sichtweisen in einer Gruppe kollektiv dargestellt und verhandelt werden. Ebenso können Ideen, Meinungen und Sichtweisen, die in Gruppendiskussionen artikuliert wurden, aus einer individuellen Perspektive vertiefend in Einzelinterviews betrachtet werden.

#### 4. Diskussionslinien, Einsatzfelder und Anwendungsvarianten der Gruppendiskussion in der Evaluation

Auch in der Programmevaluation erfahren Gruppendiskussionsverfahren seit mehreren Jahren einen Bedeutungszuwachs. Dies zeigt sich an einer Vielzahl an Evaluationsvorhaben in verschiedenen Anwendungs- und Politikfeldern, die Gruppendiskussionen als Erhebungsmethode nutzen, z.B. im Bereich der Gesundheitsforschung (vgl. Überblick in Wilkinson 1998a), um den Erfolg von Programmen

zur Gesundheitserziehung und -förderung (vgl. Basch 1987), von Rehabilitationsprogrammen (vgl. Race/Hotch/Packer 1994) und gesundheitlichen Medienkampagnen (vgl. Kitzinger 1994; Philo/Seeker/Platt/Henderson/McGlaughlin/Burnside 1994) zu bewerten und um die Kundenzufriedenheit in Gesundheitseinrichtungen zu ermitteln (vgl. Schwarz/Landis/Rowe/Janes/Pullmann 2000). Belege für ihre Nutzung finden sich in vielen anderen Feldern: in der Evaluation von Reformprozessen im Bildungs- und Sozialsystem (vgl. Gilfores/Granado 1995; Mäder/Schmidt 2005), in der Evaluation von Programmen im Bereich der Schule und Schulsozialarbeit (vgl. Nentwig-Gesemann/Bohnsack 2005; Streblov 2005) und in der Hochschulevaluation (vgl. Ernst 2006). Auch quantitativ lässt sich ein Popularitätszugewinn feststellen: So finden sich im Sozialwissenschaftlichen Forschungsinformationssystem der Gesellschaft für sozialwissenschaftliche Infrastruktureinrichtungen (SOFIS), das seit 2000 abgeschlossene und laufende Forschungsvorhaben aus der Bundesrepublik Deutschland, aus Österreich und der Schweiz dokumentiert, von 2000 bis 2003 lediglich 12 Evaluationsvorhaben, die Gruppendiskussion als Methode angeben. Von 2004 bis 2007 steigt die Zahl auf 75 Evaluationsprojekte. Auch für den Zeitraum 2008 bis 2011 liegt die Anzahl in diesem Bereich.

Betrachtet man die deutschsprachige Literatur der Sozialforschung eingehender bezüglich der Verwendung der Methode in Evaluationsvorhaben zeigen sich in den zwei deutschsprachigen Monographien über die Gruppendiskussion keine (vgl. Loos/Schäffer 2001) bis knappe Ausführungen (vgl. Lamnek 2005). Siegfried Lamnek weist in einem Abschnitt kurz auf die Anwendungsmöglichkeit in der Programmevaluation hin, wobei er der Methode in diesem Feld ein großes Leistungspotenzial zumisst. In der zweiten Auflage seines Lehrbuchs führt er aus, dass „standardisierte Einzelmethoden nur begrenzt gültige Evaluationen liefern. Nur eine offene, flexible, mithin qualitative Methode erscheint hierfür in besonderer Weise geeignet. Die Gruppendiskussion bietet sich hier vorzüglich an“ (Lamnek 2005: 77). Seiner Meinung nach liegt die Stärke des Verfahrens in der Evaluation darin, dass mittels der Gruppendiskussion Bewertungen der Maßnahmen und ihrer Folgen aus der Perspektive verschiedener Beteiligter und Betroffener eingeholt werden können. Die Überprüfung der Zielerreichung eines Programms – die Resultatsfeststellung – ließe sich auch mit anderen – standardisierten – Instrumenten bewerkstelligen.

Im Bereich der qualitativen Evaluationsforschung nimmt die Methode derzeit an Bedeutung zu, was sich an der Anzahl der Beiträge, die sich auf Forschungen mit Gruppendiskussionen stützen, im einschlägigen Handbuch von Uwe Flick (2006b) „Qualitative Evaluationsforschung“ ablesen lässt (vgl. Bohnsack 2006; Ernst 2006; Mensching 2006; Nentwig-Gesemann 2006). Dies veranlasst sogar den Herausgeber in seinem eigenen Artikel „Interviews in der qualitativen Evaluationsforschung“, in welchem er Interviews als die zentrale Methode in der qualitativen Evaluationsforschung bezeichnet, zu der Bemerkung: „Wenn man sich die Beiträge in diesem Band anschaut, könnte dagegen der Eindruck entstehen, in der qualitativen Evaluationsforschung würden Gruppenverfahren dominieren“ (Flick 2006a: 215). In seinem Aufsatz plädiert er dafür, „die existierenden qualitativen



Methoden daraufhin zu beleuchten, inwieweit sie im Kontext qualitativer Evaluationsforschung – und zwar im Kontext angepasst und trotzdem systematisch – eingesetzt werden können.“ (ebd.: 214). Für die Gruppendiskussion ist eine weitere Kontextualisierung vor dem Hintergrund evaluationstheoretischer und -praktischer Gesichtspunkte noch zu leisten.

Eine Forschergruppe um Ralf Bohnsack und Iris Nentwig-Gesemann (vgl. Bohnsack 2006; Bohnsack/Nentwig-Gesemann 2010a und 2010b; Nentwig-Gesemann 2006) setzt sich derzeit für eine Übertragung des in der Grundlagenforschung entwickelten Modells einer selbstläufigen Gruppendiskussion in Evaluationsprojekte ein (siehe ‚Gruppendiskussion zur Ermittlung kollektiver Orientierungsmuster‘ in Kap. 2). Erstmals wurde diese Variante der Gruppendiskussion in einer Evaluationsstudie zur Erfassung von Einstellungs- und Verhaltensänderungen Jugendlicher, die an einem Schulprojekt zur Peer-Mediation teilgenommen haben, erprobt (vgl. Nentwig-Gesemann/Bohnsack 2005). Seitdem kam die Gruppendiskussion in Evaluationsprojekten unterschiedlicher Felder wie der Schulentwicklung, Jugendarbeit, -förderung sowie in der Organisations- und Personalentwicklung zur Anwendung. Die Evaluationen der Forschergruppe zeichnen sich durch spezielle methodologische und theoretische Prämissen aus, die unter dem Label ‚dokumentarische Evaluationsforschung‘ zusammengefasst werden. Der Gruppendiskussion als Erhebungsmethode kommt in dieser Evaluationsrichtung ein zentraler Stellenwert zu: Von 13 Praxisbeiträgen in ihrem einschlägigen Buch zur dokumentarischen Evaluationsforschung stützen sich neun vorrangig auf die Ergebnisse von Gruppendiskussionen (vgl. Bohnsack/Nentwig-Gesemann 2010b). Insbesondere der methodologische Einführungsartikel von Iris Nentwig-Gesemann (2010) beinhaltet Überlegungen zur Anpassung der Gruppendiskussion an den Evaluationskontext. Hierin führt sie aus, dass Gruppendiskussionen in der dokumentarischen Evaluationsforschung aus einem selbstläufigen und einem evaluativen Teil bestehen. Während mittels der narrativen Passagen das „implizite, atheoretische, handlungspraktische oder konjunktive Wissen“ (Nentwig-Gesemann 2010: 65) erhoben wird, sollen die bewertenden Aussagen das „reflexive, theoretische oder auch kommunikative Wissen“ (ebd.) zugänglich machen. Für diesen zweiten bewertenden Teil der Gruppendiskussion, der beim Einsatz der Gruppendiskussion in der Grundlagenforschung bislang nachrangig war, werden zusätzliche Hinweise zu Aufbau und Gesprächsführung gegeben. Zudem wird in Analogie zu den Fokussierungsmetaphern im selbstläufigen Gruppendiskussionsteil für interaktiv dichte und emotional aufgeladene Gesprächspassagen im zweiten, evaluativen Teil der Gruppendiskussion der Begriff „Sequenzen fokussierter Argumentation“ (ebd.: 67) eingeführt. An dieser Variante der Gruppendiskussion ist für die Evaluation besonders interessant, dass sie geeignet zu sein scheint, Erkenntnisse zu teils impliziten Handlungsstrategien professioneller Akteure sowie zu Veränderungen in den Einstellungen und im Verhalten der Zielgruppen, die durch die Zielgruppe selbst eher nicht zu explizieren sind, bereitzustellen.

In angelsächsischen Evaluationszeitschriften findet sich eine Reihe von Aufsätzen, welche die Leistungspotenziale der Methode feldspezifisch im Rahmen der Evaluation von Gesundheitsprogrammen diskutieren (vgl. Richter/Bottenberg/

Roberto 1991) bzw. methodische Hinweise zur Anlage von Gruppendiskussionen mit speziellen Zielgruppen wie Kindern (vgl. Gibson 2007; Hoppe/Wells/Morrison/Gillmore/Wilsdon 1995) oder hörgeschädigten Personen (vgl. Balch/Mertens 1999) geben. Einzelne Aufsätze erörtern spezifische Nutzungsmöglichkeiten im Evaluationsprozess – z.B. als Methode zur Instrumentenentwicklung (vgl. Allen/Boley Cruz/Leonard/Unger 2010; Perreault/Renaud/Bourassa/Beauchesne/Mpiana/Bernier 2010), zur Bedarfsermittlung, Programmplanung und -evaluation (vgl. Buttram 1990; Wyatt/Krauskopf/Davidson 2008) oder zur Ermittlung von Evaluationskriterien (vgl. Moro/Cassibba/Constantini 2007). Einige wenige Artikel beschäftigen sich mit der Auswertung von Gruppendiskussionen in Evaluationsprojekten (vgl. Massey 2011; Wiggins 2004). Teilweise wird die Bedeutung von Gruppendiskussionen für verschiedene Evaluationsaufgaben wie die Einbeziehung von Stakeholdern (vgl. Abma 2006), die Exploration ihrer Werte und Interessen (vgl. Orr 2010) oder die Rekonstruktion ihrer Programmtheorien (vgl. Hansen Balle/Vedung 2010) benannt.

In den englischsprachigen Lehrbüchern zur Evaluation liegt der Schwerpunkt weniger auf der vertieften Betrachtung einzelner Methoden als vielmehr auf der Vorstellung von Strategien und Evaluationskonzepten in Anlehnung an die favorisierten Evaluationsmodelle der Autorinnen und Autoren (vgl. Rossi/Freeman/Lipsey 2003; Stake 2004). Außerdem ist die Tiefe der Darstellung einzelner Methoden in Standardwerken aufgrund ihres beabsichtigten Überblickscharakters begrenzt. In seinem Standardwerk „Qualitative Evaluation und Research Methods“ skizziert Michael Quinn Patton (1990) auf zwei Seiten die Charakteristika von Focus Group Interviewing und schreibt der Methode breite Einsatzmöglichkeiten in der Programmevaluation zu. „Focus Group Interviews can be used at any point in the evaluation process. [...] In short, focus groups can be used for a full range of evaluation purposes.“ (Patton 1990: 336). In der Neuauflage (2002) beschreibt er die Focus Group gemäß dem veränderten Titel „Qualitative Research and Evaluation Methods“ stärker als qualitative Forschungsmethode und weniger als Evaluationsmethode. Immerhin widmet er der Focus Group diesmal vier Seiten.

Einschlägige Lehrbücher zur Gruppendiskussion verweisen meist allenfalls kursorisch auf die Einsatzmöglichkeiten in der Programmevaluation. Morgan und Krüger (1998) nennen in Band 1 ihres sechsbändigen Werks „Focus Group Kit“ vier Einsatzfelder der Fokusgruppe in der Evaluationsforschung: Bedarfsanalyse, Programmentwicklung, Prozessevaluation, Ergebnisevaluation. Unter den Fallbeschreibungen finden sich Beispiele für den Einsatz von Fokusgruppen in einer Bedarfsanalyse und einer Ergebnisevaluation. Neuere Lehrbücher äußern sich häufig nicht zu Anwendungsmöglichkeiten in der Evaluation (vgl. Hennink 2007; Liamputtong 2011).

Über die Kombinationsmöglichkeiten der Gruppendiskussion mit anderen Erhebungen bzw. Methoden sowie ihre Nutzungsmöglichkeiten in Evaluationsdesigns ist aus der Literatur somit sehr wenig zu erfahren. Die Ausführungen bleiben hier eher auf einem oberflächlichen Niveau. Aus einer eigenen Analyse von rund 25 deutschsprachigen Evaluationsberichten, die die Autorin im Rahmen ihres laufenden Dissertationsvorhabens durchgeführt hat, werden im Folgenden erste

Hinweise zu den Verwendungen von Gruppendiskussionen in Evaluationen abgeleitet. So ist festzustellen, dass Gruppendiskussionen eher nicht als einzige Datenerhebungsmethode in Evaluationsprojekten zur Anwendung kommen. Auch eine Nutzung der Methode als nachgeordnete Ergänzung zu anderen Erhebungen („Supplement“) ist weniger anzutreffen. In der Regel werden Gruppendiskussionen als eine Erhebung neben anderen in einem multiperspektivischen und multimethodischen Evaluationsdesign eingesetzt. Begleitende Erhebungen können sowohl standardisierte Erhebungen wie schriftliche oder Online-Befragungen als auch weitere qualitative Zugänge wie telefonische oder Face-to-Face-Interviews und Projektbesuche sein. In den verschiedenen Erhebungen werden teilweise dieselben Fragestellungen aus unterschiedlichen Perspektiven (wie Programmverantwortliche, -mitarbeitende oder -zielgruppen) untersucht. Ebenso verbreitet ist die Verfolgung von je nach Untersuchungsgruppe unterschiedlichen Fragestellungen. Hierbei handelt es sich um Themenbereiche, zu denen eine bestimmte Beteiligtegruppe in besonderer Weise Auskunft geben kann (wie Projektmitarbeitende über die Projektumsetzung oder Projektverantwortliche über die strategische Ausrichtung des Projekts). Die Designs der Gruppendiskussionen in den näher betrachteten Evaluationsprojekten bestehen aus einer Gruppendiskussion mit ausgewählten Personen bis hin zu mehreren Gruppendiskussionen (in der Regel zwischen 2 und 8). Es finden sich u.a. Projekte, bei denen mit denselben Personen zu verschiedenen Zeitpunkten im Programmverlauf Gruppendiskussionen durchgeführt werden, um Veränderungen in Einstellungen und Verhalten der Personen zu ermitteln. Häufiger noch werden mehrere Gruppendiskussionen mit einer ähnlichen Teilnehmer(innen)struktur wie Mitarbeitende verschiedener Standorte oder Zielgruppenmitglieder verschiedener Teilprojekte umgesetzt. Diese Vorgehensweise erlaubt es, die Heterogenität der Gruppen in den Erhebungen besser abzubilden. Vorzufinden ist zudem die Durchführung von Gruppendiskussionen mit unterschiedlicher Teilnehmer(innen)zusammensetzung (beispielsweise mehrere Gruppendiskussionen mit jeweils Projektmitarbeitenden und Projektzielgruppen), um eine multiperspektivische Betrachtung von Themen, Entwicklungen und Meinungen zu ermöglichen.

## 5. Besonderheiten von Evaluation

Die im Rahmen von Evaluation zu bewältigenden Herausforderungen und Aufgaben unterscheiden sich von denen in der Sozialforschung in einigen Punkten. Dies soll im Folgenden kurz veranschaulicht werden, um hieran den Nutzen von Gruppendiskussion in der Evaluation zu verdeutlichen.

Auch wenn sich die Profession uneinig ist, ob Evaluation als ein Teilbereich empirischer Forschung, welcher keine eigene Disziplin darstellt (vgl. Bortz/Döring 2002), bzw. als eine besondere Form sozialwissenschaftlicher Forschung (vgl. Haubrich/Holthausen/Struhkamp 2005; Lüders 2006) zu verstehen ist, oder ihr ein eigenständiger Status zuzusprechen ist (vgl. Beywl 1988; Wesnierski 2003), ist es unstrittig, dass Evaluationen gegenüber der sozialwissenschaftlichen Grundlagen-

forschung und der angewandten Forschung Besonderheiten aufweisen (vgl. Bohnsack/Nentwig-Gesemann 2006; Kardoff 2005 und 2006; Stockmann 2000; Wottawa/Thierau 2003). Im Kern lassen sich drei zentrale Merkmale von Evaluationen benennen, die spezifische Erkenntnisinteressen verdeutlichen und darauf zugeschnittene eigene methodische Zugänge erfordern.

### *Praxis- und Nutzenorientierung*

Ein erstes Unterscheidungskriterium ist die Erwartung, dass Evaluationen für gesellschaftliche Akteure nützliche und verwertbare Erkenntnisse bereitstellen. Während Forschung – zugespißt formuliert – in erster Linie darauf abzielt, kumulatives, raum-zeitlich überdauerndes Wissen zu Gegenstandsfeldern zu liefern, sollen Evaluationen meist „pragmatische und plausible Antworten auf praktisch-politische Fragen“ (Kardoff 2006: 69) liefern. Mit dieser Ausrichtung geht einher, dass die Adressatinnen und Adressaten der Ergebnisse von Evaluationen nicht primär Wissenschaftlerinnen und Wissenschaftler, sondern vorrangig institutionelle und soziale Akteure sind, die die Ergebnisse nach Maßstäben wie Relevanz, Aktualität und erfahrungsgestützte Gültigkeit bewerten (vgl. Hammersley 2003). Evaluation ist damit analog zur Auftragsforschung zu verstehen, die sich an den Fragestellungen der Auftraggebenden und ggf. weiterer Beteiligter orientiert. In der Evaluationspraxis sind damit notwendigerweise häufigere Absprachen über bestehende Informationsbedarfe, anvisierte Nutzungszeitpunkte sowie geeignete und adressatengerechte Formen der Ergebnispräsentation verbunden.

Michael Q. Patton (1998) hat herausgestellt, dass nicht nur die Ergebnisse von Evaluation geeignet sind, einen Zugewinn an Erkenntnis und Qualität zu ermöglichen, sondern auch der Prozess der Evaluation einen positiven Nutzen für das Programm, die Auftraggebenden, die Beteiligten und Betroffenen hervorbringen kann. Ein Prozessnutzen entsteht beispielsweise, wenn Zielgruppenmitglieder oder Projektakteure durch ihre Einbeziehung in die Evaluation ihre Handlungsfähigkeiten erweitern. In vielen Evaluationsprojekten wird daher der Austausch zwischen Mitgliedern der Zielgruppe oder anderen Beteiligtegruppen absichtsvoll ermöglicht, so dass die Personen alternative Handlungsstrategien kennenlernen und ihre Problemlösungskompetenz erhöhen.

Auch wenn von einzelnen Vertretern der Profession (vgl. Lüders 2006) die „direkte Verwendungsorientierung“ (Kromrey 2003: 93) von Evaluation als konstitutives Merkmal mit dem Hinweis auf die nicht kalkulierbaren, selektiven und strategischen Nutzungsformen der Praxis, die ihrer eigenen Logik folgen, abgelehnt wird, und auf die Brüchigkeit der Abgrenzung zur sogenannten Grundlagenforschung – aufgrund ihrer zunehmenden Kopplung mit den sie umgebenden Umwelten – hingewiesen wird, kann weiterhin von einem graduell höherem Maß an Praxisorientierung in Evaluationen im Vergleich zur (Grundlagen-)Forschung ausgegangen werden. Dies erfordert verstärkt methodische Zugänge, die eine Einbeziehung der Programmbeteiligten und -betroffenen in die Gestaltung von Evaluationen erlauben.

### *Bewertung*

Wie bereits der Terminus Evaluation (*to evaluate – bewerten*) nahelegt, wird von Evaluationen erwartet, dass sie Gegenstände der sozialen Wirklichkeit mittels empirisch gewonnener Informationen bewerten. In nahezu allen Definitionen wird das Bewerten und Urteilen als konstitutive Evaluationsaufgabe – mehr oder weniger explizit – benannt. So versteht die Gesellschaft für Evaluation unter Evaluation „die systematische Untersuchung des Nutzens oder Wert eines Gegenstands“ (DeGEval 2002).

Viele Evaluierende sehen in dem mit Evaluation verbundenen Bewertungsanspruch den zentralen Unterschied zur sozialwissenschaftlichen Forschung. „Evaluation unterscheidet sich von sonstigen Formen, meist als ‚wertfrei‘ verstandener sozialwissenschaftlicher Forschung insbesondere durch ihren Bewertungsanspruch“ (vgl. Haubrich/Holthusen/Struhkamp 2005: 1). In seinem Beitrag „Qualitative Evaluationsforschung – was heißt hier Forschung?“ führt Lüders aus, dass von Evaluationsforschung „eben nicht nur (die) ordentliche empirische Analyse und Rekonstruktionen sozialer Sachverhalte, sondern eben auch eine besonders verlässliche Qualität der Bewertung dessen, was da beobachtet und analysiert worden ist“ (Lüders 2006: 50) erwartet wird. Für dieses seiner Meinung nach vorrangige Alleinstellungsmerkmal von Evaluation konstatiert er unbedingten Forschungs- und Handlungsbedarf, da es bislang weder methodologisch noch methodisch geklärt sei, wie Bewertungen systematisch erfolgen können.

So wird der Bewertungsprozess in Evaluationen sehr unterschiedlich organisiert: Wer legt die Bewertungsmaßstäbe fest? Auf der Grundlage welcher Wertpositionen erfolgt die Bewertung? Ist der Bewertungsakt als zentrale Aufgabe der Evaluation angelegt oder eher nachrangig? Beywl (2006) zeigt auf, wie unterschiedlich die bestehenden evaluationstheoretischen Modelle die Bewertungsakte anlegen. Die Spannbreite reicht von Modellen, die den Bewertungsprozess außerhalb der Evaluation verorten (wertedistanziert), und solchen, die aus einer unparteiischen Position heraus intensiv mit Wertgemeinschaften und Wertkonflikten arbeiten (werterelativiert). Andere sehen Aushandlungsspielräume für soziale Werte vor und legen über systematische Priorisierungen Wertekonsense offen (wertepriorisierende Modelle). Wertepositionierte Ansätze privilegieren Werte und Interessen von Programmmitarbeitenden oder auch Zielgruppen. Je nachdem, wie die Bewertungsprozesse in der Praxis organisiert werden, sind in unterschiedlichem Umfang Aushandlungsprozesse mit Auftraggebenden und den verschiedenen Stakeholdergruppen<sup>2</sup> zu planen und umzusetzen.

---

2 Stakeholder sind Personen, Personengruppen oder Organisationen, die, je nachdem, wie ein Programm konzipiert und umgesetzt wird, etwas zu verlieren oder zu gewinnen haben: Für sie steht etwas auf dem Spiel, sie halten Einsätze (engl. *stakes*). Es können aktiv am Evaluationsgegenstand Beteiligte von passiv Betroffenen unterschieden werden. In: Eval-Wiki: Glossar der Evaluation, Univation, Institut für Evaluation, Stand 30.03.2011. Verfügbar unter: [http://www.eval-wiki.org/glossar/Stakeholder\\_%28eines\\_Programms%29](http://www.eval-wiki.org/glossar/Stakeholder_%28eines_Programms%29) [20.01.2013].

### *Dialogorientierung*

Gewissermaßen als Querschnittsaufgabe ergibt sich – insbesondere bedingt durch die Orientierung an den Informationsbedarfen der gesellschaftlichen Akteure und die interaktive, organisationsförmige Beschaffenheit der Evaluationsgegenstände – die Notwendigkeit zu einem Dialog mit den relevanten Beteiligten wie Auftraggebenden, Programmmitarbeitenden oder auch anderen wichtigen Gruppen wie Expert(inn)en, der Fachöffentlichkeit und der Zielgruppe. Bohnsack (2006) bezeichnet dies als den kommunikativen und interaktiven Charakter, den *Gesprächscharakter*, der Evaluation. Zum einen versteht er darunter die Analyse der Gespräche relevanter Beteiligungsgruppen (z.B. basierend auf Beobachtungsprotokollen oder elektronischen Aufzeichnungen) mit dem Ziel, die kollektiven Wissensbestände und Orientierungsmuster der Akteure zu rekonstruieren. Zum anderen ist es aus seiner Sicht eine wesentliche Aufgabe der Evaluation, einen Diskurs zwischen den unterschiedlichen Gruppen in Gang zu bringen.

Insbesondere dieser zweite Aspekt stellt eine Besonderheit von Evaluation im Vergleich zu anderen Formen sozialwissenschaftlicher Forschung dar. So sind in allen Phasen oder Leistungsschritten der Evaluation kommunikative und interaktive Leistungen zu erbringen. In der Phase der Gegenstandsbestimmung geht es darum, gemeinsam mit den Stakeholdergruppen zu einem geteilten Programmverständnis zu gelangen und die Programmlogik zu explizieren. Weiterhin gilt es zu klären, welche Zwecke die gesellschaftlichen Akteure mit der Evaluation verfolgen, also wie die Evaluationsergebnisse genutzt werden sollen. Meist müssen zudem nach der Auftragserteilung im Austausch mit relevanten Beteiligten die Fragestellungen konkretisiert und eventuell priorisiert werden. Auch dies erfordert dialogische Verfahren wie moderierte Workshops. Aber auch in der Phase der Informationsgewinnung, der Datenerhebung und -auswertung, sind kommunikative Leistungen zu erbringen. So kommt es bei einer Vielzahl von empirischen Zugängen wie Interviews oder Gruppendiskussionen zu einer Interaktion zwischen den Evaluierenden und den Personen oder Gruppen, die Daten zur Ermittlung der sozialen, politischen oder pädagogischen Praxis liefern sollen. Auch die Vermittlung der Ergebnisse ist durch einen kommunikativen Prozess der Diskussion zwischen den Evaluierenden und den jeweiligen Projektakteuren gekennzeichnet. Dieser ist durch die Evaluation vorzubereiten und einzuleiten, wengleich dies sicherlich je nach Budget, Evaluationskonzept und Interessen der Beteiligungsgruppen unterschiedlich umfangreich erfolgt.

Aus diesen spezifischen Aufgaben und Leistungen, denen in der evaluativen Praxis zu entsprechen ist, lassen sich folgende Schlussfolgerungen ziehen:

Ergänzend zur systematischen Informationsgewinnung sind in Evaluationen zusätzliche Aufgaben wie die Klärung des Gegenstands, die Bewertung der beobachteten Praxis und die Vermittlung der Ergebnisse zu erfüllen (vgl. Beywl 2006). Diese Aufgaben erfordern in einem hohen Maße kommunikative und interaktive Leistungen. Gruppenmethoden, welche die Gruppe als Setting zur Erhebung von Daten und zur Gestaltung des Evaluationsprozesses nutzen, sind in besonderem Maße geeignet, hierzu einen Beitrag zu leisten.

## 6. Leistungspotenziale und Grenzen

In Gruppendiskussionen können Daten erzeugt werden, die andere Methoden gar nicht, nicht in dieser Form oder nur mit erheblich größerem Aufwand bereitstellen können. Über viele allgemeine Vorteile der Methode besteht weitgehend Einigkeit. Gruppendiskussionen sind demnach vor allem produktiv, wenn Sichtweisen, Einstellungen, Geschichten und Selbstrepräsentationen in einem gegebenen kulturellen und kollektiven Kontext erhoben werden sollen. Ihre Stärke liegt darin, zu ermitteln, wie Meinungen und Sichtweisen konstruiert sind und artikuliert werden (vgl. Flick/von Kardoff/Steinke 2005; Kitzinger 2005; Lamnek 2005; Liamputtong 2011; Wilkinson 2004).

In der Programmevaluation hat die Gruppendiskussion gegenüber anderen Methoden spezifische Vorzüge, was eine weitere Betrachtung und Fundierung der Methode als Evaluationsmethode nahelegt:

Gruppendiskussionen sind geeignet, emotionale Hintergründe von Äußerungen sowie verborgene Annahmen und Meinungen aufzudecken. Myers und Macnaghten (1999) belegen mittels konversationsanalytischer Betrachtungen der Gruppeninteraktion, dass Gruppenmitglieder Standpunkte und Haltungen in Relation zu anderen Gruppenmitgliedern wie Teilnehmenden und Moderierenden einnehmen, da sie im Gruppengespräch zu Positionierungen herausgefordert werden. Demzufolge sind die ausgelösten Stellungnahmen keine von vorneherein festgelegten Attribute, welche ‚nur‘ abgefragt zu werden brauchen. In anderen Konversationssettings können solche Meinungen daher eher implizit bleiben. So können beispielsweise kritische Nachfragen von Programmmitarbeitenden oder -verantwortlichen zur Umsetzung oder zu den Entwicklungen eines Programms an einem anderen Standort zur vertieften Darstellung von Einschätzungen und Zusammenhängen anregen (vgl. Belzile/Öberg 2012).

Infolge des interaktiven Charakters werden Daten auf hohem Artikulationsniveau durch sich ergänzende, divergierende und ausführliche Beschreibungen der Teilnehmenden zu einem spezifischen Ausschnitt produziert (vgl. Wilkinson 1998a). Die in Gruppendiskussion auftretenden Kommunikationsformen, wie Anekdoten, Erzählungen und Witze, offenbaren erweiterte Verständnismöglichkeiten (vgl. Kitzinger 1994; Wilkinson/Rees/Knight 2007). So kann der Austausch von episodischem und erfahrungsnahem Wissen, z.B. von Zielgruppen eines Programms wie Flüchtlingen in Beratungsprojekten, zusätzliche Hinweise über Erfahrungen und Wahrnehmungen innerhalb eines Programms bieten.

Die Teilnehmenden regen sich gegenseitig zu freimütigen und offenen Beiträgen an – nach der Devise: „Wir sind ja unter uns ...!“ Entgegen der Annahme, dass insbesondere durch individuelle Interviews vertrauliche und sensible Daten erhoben werden können, hat sich herausgestellt, dass gerade die Gruppe – bestehend aus einander bekannten Personen bzw. Personen mit einem ähnlichen Erfahrungshintergrund – katalytisch in der Diskussion sensibler Themen wirkt. Gerade auch bei stigmatisierten oder sozial benachteiligten Gruppen (wie Sozialhilfebeziehenden) stützt die Gruppensituation die Einzelnen. Diese Personen sind häufig eher bereit,

an einer Gruppenerhebung als an einem Einzelinterview teilzunehmen. Auch ist in solchen Fällen gegenüber Personen in ähnlichen Lebenslagen von einer größeren Offenheit auszugehen (vgl. Madriz 1998). Gruppendiskussionen können somit auch in Evaluationssettings, in denen mit benachteiligten Gruppen gearbeitet wird, passende methodische Zugänge darstellen.

Gruppendiskussionen initiieren Lernprozesse zwischen den Beteiligten (vgl. Wibeck/Dahlgren/Öberg 2007). Sie ermöglichen die Entwicklung eines kollektiven Verständnisses eines geteilten Problems und bieten einen Rahmen für die Entwicklung von Lösungsansätzen. Durch den Austausch von Erfahrungen, Einstellungen und Interpretationen wird eine Meinungsbildung in der Gruppe auch zu komplexen Fragestellungen möglich. Aus denselben Erwägungen heraus wird die Fokusgruppe vermehrt in der Aktionsforschung – welche zentrale Anliegen mit der beteiligten-orientierten Programmevaluation teilt – zur Förderung von politischen und sozialen Veränderungen eingesetzt (vgl. Barbour/Kitzinger 1999; Kamberelis/Dimitriadis 2008 und 2011; Wilkinson 1998a).

Die Teilnahme an Gruppendiskussionen hat positive Kommunikationserfahrungen zur Folge, wodurch mittelfristig auch das evaluierte Programm profitiert. Der Austausch programmspezifischer Argumente und Perspektiven führt bei den Teilnehmenden meist zu einer größeren Identifikation mit den Programmenthemen. Die Teilnehmenden interessieren sich in der Folge stärker für das Programm und übernehmen eher Verantwortung für die Weiterentwicklung und Umsetzung des Programms.

Teilnehmende fühlen sich in Gruppendiskussionen als gleichberechtigte Diskussionspartner(innen) wertgeschätzt. Dies unterstützt den Aufbau einer vertrauensvollen Arbeitsbeziehung zwischen Evaluierenden und Programmbeteiligten. Nach Wilkinson macht der geringere Einfluss der Forschenden auf die Erhebungssituation die Gruppendiskussion zu einer eher egalitären Methode, wodurch sie insbesondere auch für die feministische Forschung als geeignete Methode propagiert wird (vgl. Wilkinson 1999a, 1999b). Diese weniger hierarchische Kommunikationssituation verleiht den Sichtweisen und Standpunkten der Teilnehmenden ein stärkeres Gewicht (vgl. Madriz 2003; Morgan 1988). Die Teilnehmenden haben eine größere Chance, die Diskussionsinhalte mitzubestimmen und so Themen zu diskutieren, die für sie selbst von Bedeutung sind (vgl. Wilkinson 1999b).

Weniger geeignet sind Gruppendiskussionen bei den folgenden Anliegen und Rahmenbedingungen (vgl. Kitzinger 2005; Smithson 2008): Wenn die Exploration individueller Biographien und Erfahrungen im Vordergrund steht, sind Einzelinterviews zu präferieren. Für die Analyse von sozialen Rollen sowie formalen und informellen Prozessen in Organisationen und Projekten sind eher Beobachtungen zu empfehlen. Sollen in Evaluationen vergleichbare Informationen zu lokalen Projektcharakteristika erfragt werden (beispielsweise zur personellen und finanziellen Ausstattung von Projekten, ihrer Reichweite oder ihrem Umsetzungsgrad), bieten sich standardisierte Befragungen an. Gruppendiskussionen sollten weiterhin nicht eingesetzt werden, wenn die Verbreitung bestimmter Meinungen in einer größeren Population untersucht werden soll. Hierfür sind standardisierte Erhebungen größeren Maßstabs erforderlich. Über standardisierte Befragungen können zum Beispiel



Diskrepanzen zwischen den Zielen und Zielerreichungen innerhalb eines Projekts deutlich werden. Gruppendiskussionen bieten eher einen Rahmen, um Diskrepanzen dieser Art in der Gruppe zu erklären und zu diskutieren.

Wichtige Voraussetzungen für erfolgreiche Gruppendiskussionen sind, dass die Teilnehmenden an dem Diskussionsthema interessiert sind und etwas zu der Diskussion beizutragen haben. Sind ihnen die Themen fremd bzw. werden sie als unwichtig eingestuft, entsteht nicht die anzustrebende kreative und angenehme Atmosphäre unter den Beteiligten der Gruppendiskussionen. Ein weiteres Hemmnis für den Diskussionsverlauf stellen Spannungen, Abhängigkeiten und Wertkonflikte zwischen den Teilnehmenden dar. Insofern ist im Vorfeld zu antizipieren, ob die eingeladenen Personen die anvisierten Themen möglichst spannungs- und sanktionsfrei miteinander diskutieren können. Demgemäß schlägt Morgan (1997: 17) vor, „to ask how actively and easily the participants would discuss the topic of interest“.

## 7. Formen der Gruppendiskussion

Auf einem Kontinuum bewegen sich die in der evaluativen Praxis eingesetzten Verfahren zwischen zwei Polen: Auf der einen Seite befinden sich strukturierte Varianten, gekennzeichnet durch direktives Moderatorverhalten mit kürzeren wechselseitigen Gesprächszyklen zwischen den Teilnehmenden und einem engeren Themenfokus; auf der anderen Seite stehen gering gesteuerte, selbstläufige Diskussionsformate mit hohen dialogischen Anteilen auf Seiten der Teilnehmenden, entlang selbst gesetzter Themenschwerpunkte. Stellvertretend für die selbstläufige Form ist das Modell zur Erfassung kollektiver Orientierungsmuster nach Bohnsack (vgl. Bohnsack/Przyborski/Schäffer 2006) zu nennen, während die strukturierte Variante insbesondere durch Morgan und Krueger (vgl. Krueger 1994; Morgan 1993) aus der Marktforschung in die sozialwissenschaftliche Praxis eingeführt wurde. Die Unterschiede zwischen den beiden Formen werden idealtypisch in der nachfolgenden Tabelle abgebildet (vgl. Tabelle 1).

*Tabelle 1:* Formen der Gruppendiskussion

	Strukturierte Form	Selbstläufige Form
<i>Themenbreite</i>	eng	breit
<i>Zusammensetzung</i>	künstlich („Fremde“) homogen	real („Clique“) heterogen
<i>Setting</i>	formal	natürlich
<i>Diskussionsgestaltung</i>	ausgearbeitete verbale/mediale Impulse	keine bzw. grobe Themenvorgabe
<i>Moderationsverhalten</i>	leitend	zurückhaltend
<i>Datenart</i>	begrifflich-explizit	narrativ-implizit

Aufgrund der ausgeprägten Praxisorientierung sowie der geforderten Gegenstandsangemessenheit der Methoden ist davon auszugehen, dass sich in der Evaluationspraxis neben den idealtypischen selbstläufigen bzw. strukturierten Varianten des Verfahrens Mischformen bzw. dem jeweiligen Evaluationskontext angepasste Varianten des Verfahrens herausgebildet haben. Bislang stehen Überblicke über die Ausdifferenzierung von Varianten in der Evaluationspraxis noch nicht zur Verfügung. Über tatsächliche Praxisformen kann daher an dieser Stelle nicht berichtet werden. Es soll allerdings in einem ersten Schritt für die Variationsbreite sensibilisiert werden, indem zentrale Gestaltungsparameter vorgestellt werden, über die in der Praxis entschieden werden muss:

- Stellung der Gruppendiskussion im Evaluationsdesign (eigenständig, gleichwertig, nachgeordnet): Festzulegen ist, ob Gruppendiskussionen in den Evaluationen die einzige oder dominante Datenerhebungsmethode darstellen, ob sie als gleichwertige Datenerhebungsmethode in einer multimethodischen Untersuchung fungieren oder als ergänzende Untersuchung in Verbindung mit anderen dominierenden Methoden auftreten.
- Funktion der Gruppendiskussion im Evaluationszyklus (Gegenstandsbestimmung, Informationsgewinnung, Ergebnisvermittlung): In Anlehnung an die durch Friedrichs (1990) eingeführten Forschungsphasen *Entdeckungs-*, *Begründungs-* und *Verwertungszusammenhang* kann auch der Evaluationsprozess in drei Teile gegliedert werden (vgl. Beywl/Kehr/Mäder/Niestroj 2007). In der Phase der Gegenstandsbestimmung werden gemeinsam mit den Programmteilnehmenden die Evaluationszwecke, Fragestellungen und Bewertungskriterien geklärt. Im Anschluss wird das Design für die Evaluation entwickelt. Daten werden erhoben und ausgewertet (‘Informationsgewinnung’). Die Phase der Ergebnisvermittlung zielt darauf, die Nutzung der Ergebnisse vorzubereiten und einzuleiten. In allen drei Phasen des Evaluationsprozesses können Gruppendiskussionen aussagekräftige Daten liefern.
- Gruppendiskussion zur Unterstützung verschiedener Evaluationsfunktionen (Bedarfsermittlung/Zielklärung, Prozessoptimierung, Wirkungserfassung): Gruppendiskussionen dienen in Evaluationsprojekten insbesondere zur Bedarfsermittlung und Zielklärung von Programmen, zur Prozessoptimierung und zur Ermittlung der Wirkungen von Programmen. Im Folgenden werden diese drei Funktionen von Gruppendiskussionen in Evaluationen näher ausgeführt (vgl. Mäder 2005). Bei der Bedarfsermittlung/Zielklärung werden Fokusgruppen genutzt, um Wünsche, Anliegen und Erwartungen (potenzieller) Nutzerinnen und Nutzer von Projekten und Programmen zu ermitteln. Im US-amerikanischen Raum wurden Gruppendiskussionen vielfach von staatlichen Behörden und Non-Profit-Organisationen zur Bedarfsanalyse eingesetzt (beispielsweise um zielgruppenspezifische Informationen für Stadtplanung und -entwicklung zu erhalten) (vgl. Hennink 2007; Morgan 1998). Fokusgruppen bieten zudem einen Rahmen, um Ziele/Bewertungskriterien, Nutzenerwartungen und Risiken zu thematisieren, die mit einem Programm seitens der Beteiligten verbunden sind (vgl. Moro/Cassibba/Constantini 2007). Prozessoptimierungen unterstützen Fokusgruppen, indem sie ein geeignetes Forum bieten,

um Stärken und Schwächen im Programmverlauf zu identifizieren. Ein Austausch von Erfahrungen und Meinungen der Beteiligten und Betroffenen untereinander erhöht die Einsicht in die Wirkungsweise des Programms und kann erste Hinweise für Anpassungen und Korrekturen liefern. Weiterhin können Fokusgruppen helfen, Erkenntnisse über bereits beobachtete oder zu erwartende Effekte, zum Beispiel aus der Perspektive von Programmmitarbeitenden zusammenzutragen. Gruppendiskussionen sind ebenfalls geeignet, die erwünschten Wirkungen in Form von Veränderungen in den Meinungen und Haltungen der Programmzielgruppen zu erfassen (vgl. Nentwig-Gesemann/Bohnsack 2005). Neben den angezielten Wirkungen lösen Programme auch nicht Vorhergesehenes aus. Fokusgruppen können auch dazu genutzt werden, nicht intendierte Nebenwirkungen – sowohl erwünschte als auch unerwünschte – zu erfassen (z.B. Diskussion von Interessenvertreter(inne)n benachteiligter Gruppen über die (Neben-)Wirkungen eines Förderprojekts).

- Design der Gruppendiskussion (Anzahl bzw. Serien): Bei der Planung der Gruppendiskussion ist zu entscheiden, wie viele Gruppendiskussionen mit welcher Teilnehmer(innen)zusammensetzung durchgeführt werden. Wenn mehrere Diskussionen mit der gleichen Teilnehmer(innen)zusammensetzung realisiert werden, wird von ‚Serien‘ gesprochen (z.B. drei Gruppendiskussionen mit weiblichen Programmnutzerinnen). Die Anzahl der Diskussionen steht mit der Frage im Zusammenhang, wann eine hinreichende Meinungsbreite und Gültigkeit der Ergebnisse erzielt wird. Wenn Perspektiven unterschiedlicher Gruppen erhoben werden sollen, sind getrennte Gruppendiskussionen, beispielsweise mit Programmbeteiligten und Programmzielgruppen, zu realisieren.
- Teilnehmende der Gruppendiskussion (Fachöffentlichkeit, Geldgeber(innen), Programmverantwortliche, Programmmitarbeitende, Programmbetroffene): In Evaluationen nehmen die Beteiligten und Betroffenen eine zentrale Rolle ein. Zum einen sind sie in unterschiedlichem Umfang in die Planung der Evaluation einbezogen. Zum anderen fungieren sie als Datengebende in den Erhebungen. Beteiligte sind Personen, die in Bezug auf den Evaluationsgegenstand Interessen verfolgen und an der Planung und Umsetzung des Programms aktiv mitwirken. Als ‚Betroffene‘ werden Personen mit wenig Einfluss bezeichnet. Hierzu zählen nicht selten die Zielgruppen des Programms (vgl. DeGeval 2002). Bei den Überlegungen zur Konzeption der Gruppendiskussion sind die Kompetenzen, Befindlichkeiten und Haltungen der einzubeziehenden Gruppe zu berücksichtigen.
- Zusammensetzung der Gruppen (natürliche vs. künstliche Gruppen): Als natürlich werden Gruppen bezeichnet, die auch im Alltag bestehen bzw. in der sozialen Wirklichkeit als Gruppe existieren (z.B. Mitarbeitende einer Organisation). Künstliche Gruppen werden je nach Forschungsinteresse gemäß einzelner oder mehrerer Kriterien zusammengestellt (z.B. Arbeitsamtsmitarbeitende, Sozialhilfebeziehende). In der Literatur wird diese Unterscheidungsdimension mit unterschiedlichen Begrifflichkeiten belegt: künstliche vs. natürliche Gruppen (vgl. Flick 1995), Ad-hoc-Gruppen vs. natürliche Gruppen (vgl. Lamnek 1998), „Strangers or Acquaintances“ (Morgan/Krueger 1998), Realgruppen (vgl. Nie-

ben 1977). In Abhängigkeit von den unterschiedlichen Anwendungsfeldern und Modellen der Gruppendiskussion herrschen teilweise eindeutige Präferenzen für natürliche bzw. künstliche Gruppen vor. In jedem Fall ist durch die Form der Gruppenzusammensetzung ein Einfluss auf die Gesprächsdynamik in der Gruppe zu erwarten. Reale Gruppen sind einander vertrauter, explizieren ggf. weniger und diskutieren innerhalb ihrer Rollenstrukturen, während künstliche Gruppen sich zunächst in der Gruppe orientieren müssen, ggf. offener sprechen können und besser von der Moderation gesteuert werden können. In Evaluationen von sozialen Programmen, in denen die Beteiligten Interessen in Bezug auf den Evaluationsgegenstand verfolgen, kommt der Gruppenzusammensetzung eine gesteigerte Bedeutung zu. So muss sichergestellt werden, dass eine Ausgewogenheit der Perspektiven erreicht wird und eine angenehme, sanktionsfreie Atmosphäre geschaffen wird, in welcher die Teilnehmenden nicht strategisch agieren. Interessant ist in diesem Zusammenhang auch, wie die Teilnehmer(innen)auswahl und -gewinnung erfolgt – ein Aspekt, dem in der Literatur weniger Bedeutung zugemessen wird. Beispielsweise können Projektmitarbeitende Teilnehmende vorschlagen. In diesem Fall ist zu prüfen, ob die vorgeschlagenen Personen den Auswahlkriterien – insofern solche explizit formuliert wurden, was insbesondere bei Gruppendiskussionen mit natürlichen Gruppen seltener anzutreffen ist – entsprechen oder ob eine Ungleichverteilung in Bezug auf Kriterien vorliegt, die zu verzerrten Ergebnissen führen kann – wie z.B. wenn alle teilnehmenden Absolvent(inn)en einer Beschäftigungsmaßnahme mit einer hohen Vermittlungsquote sind. Andere Verfahren sind z.B. die Auswahl von Personen aus einer Teilnehmer(innen)liste oder die Benennung von Personen nach dem Schneeballverfahren.

- Diskussionsgestaltung (gering bis stark strukturiert): In gering strukturierten Diskussionen werden wenige offene Fragen an die Teilnehmenden gerichtet oder es wird sogar nur ein Thema zur Diskussion gestellt. Die Teilnehmenden setzen ihre Themenschwerpunkte selbst. Die Moderation hält sich weitgehend zurück und passt ihre Fragen flexibel an die sich entwickelnden Gesprächsinhalte an. In stark strukturierten Diskussionen hingegen sind die Fragen vorab ausformuliert und werden in einer vorher festgelegten Reihenfolge gestellt. Die Themenschwerpunkte sind vorgegeben und die Moderation achtet auf ihre Einhaltung. Die Gesprächsimpulse reichen von offenen Fragen bis hin zum Einsatz von Vignetten (vgl. Bailey 2008) und zur Aufforderung zu Aktivitäten wie Bilder auswählen, anhand von vorgegebenen Skalen Bewertungen vornehmen und diskutieren, Collagen anfertigen oder ein Rollenspiel durchführen (vgl. Colucci 2007).
- Moderationsverhalten (nicht direktiv vs. direktiv): Im Wesentlichen lassen sich zwei Moderationsstile unterscheiden: Ein nicht direkter Stil ist durch eine passive, zurückhaltende Haltung gekennzeichnet, bei welchem die Moderation eher die Rolle eines Beobachtenden einnimmt. Dieser Typ ist häufig in informellen und gering strukturierten Settings anzutreffen. Im Gegensatz dazu übernehmen direkte Moderierende eine aktive Rolle in Bezug auf die Steuerung der Diskussion. Dieser Stil ist häufig in formalen und stark strukturierten Set-

tings vorzufinden, bei welchen die Moderierenden auf die Beibehaltung des Fokus der Diskussion achten.

- Setting (informell vs. formell, face-to-face vs. online): Den wohl größten Gegensatz weisen die Settings von Gruppendiskussionen auf, die einerseits im Rahmen von Feldforschungen und andererseits in der Marktforschung eingesetzt werden. Erstere ergeben sich mehr oder weniger spontan an Orten, die der interessierenden Gruppe bekannt sind und die sie selbständig aufsucht, wie am Arbeitsplatz, an der Straßenecke, in der Kneipe oder auf dem Spielplatz (vgl. Frey/Fontana 1991). Letztere finden in einem formalen, vorab arrangierten Setting statt, gewöhnlich in einem extra für diese Zwecke ausgestatteten Raum mit Einwegspiegel und qualitativ hochwertigen Ton- und Videoaufnahmegegeräten (vgl. Puchta/Wolff 2004). Die Wahl des Diskussionsorts in Evaluationsvorhaben wird voraussichtlich sowohl durch das jeweilige Erkenntnisinteresse als auch von pragmatischen Erwägungen beeinflusst. Sehr häufig finden Gruppendiskussionen beispielsweise in den Räumlichkeiten von programmteiligen Institutionen statt. Aber auch die Wahl sogenannter ‚neutraler‘ Räumlichkeiten wie ein Hotel oder auch Wertschätzung ausdrückende Orte wie ein Rathaus sind üblich. Neben Face-to-Face-Gruppendiskussionen werden zunehmend auch Varianten von Online-Gruppendiskussionen eingesetzt. Anzutreffen sind Online-Diskussionen in schriftlicher Form (Chats), als verbale Kommunikation (Telefonkonferenzen) oder als eine Kombination von visueller und verbaler Kommunikation (z.B. über Skype) (vgl. Gaiser 2008; Galloway 2011).
- Person des Moderators bzw. der Moderatorin (Evaluierende oder externe Personen): Nach Wolff und Puchta (2007) sind die Moderatorinnen und Moderatoren von Gruppendiskussionen in der Marktforschung wesentlich erfahrener und professioneller als ihre Kollegen in den Sozialwissenschaften. Wenn Gruppendiskussionen in Evaluationen durchgeführt werden, muss entschieden werden, ob die Moderation durch die Evaluierenden übernommen wird oder durch eine nicht in die Evaluation involvierte Person. In jedem Fall ist im Vorhinein zu überlegen, wie die moderierende Person von den Teilnehmenden wahrgenommen werden wird und inwieweit in dieser Konstellation eine offene Gesprächsatmosphäre hergestellt werden kann. In Evaluationen, die der Entscheidungsfindung bzw. Rechenschaftslegung dienen sollen, können Evaluierende in der Rolle der Moderation bei den Teilnehmenden Widerstände und Befangenheit auslösen.
- Auswertung (thematisch, dokumentarisch, diskursiv): Die Abschnitte zur Auswertung von Gruppendiskussionen in einschlägigen Hand- und Lehrbüchern sind in der Regel eher kurz, im Gegensatz zu ausführlichen Hinweisen und Vorschlägen zur Durchführung von Gruppendiskussionen. Dazu passt, dass in veröffentlichten Untersuchungen mit Gruppendiskussionen kaum beschrieben ist, wie die Analyse durchgeführt wurde. Am weitesten ist die thematische Analyse in Evaluationszusammenhängen verbreitet (vgl. Massey 2011; Wiggins 2004). Dies lässt sich darauf zurückführen, dass in Evaluationen mithilfe von Gruppendiskussionen meistens die Haltungen und Einstellungen von Stakeholder-Gruppen gegenüber dem Programm ermittelt werden sollen. Mit der thema-

tischen Analyse wird nach gemeinsamen Themen bzw. der thematischen Breite innerhalb einer Gruppe gesucht. Massey (2011) unterscheidet drei Datenarten, die durch die thematische Analyse gewonnen werden können. Artikulierte Daten („articulated data“) entstehen demnach durch die direkten Antworten der Teilnehmenden auf die Fragen bzw. Nachfragen der Moderation. Auf diese Datenart wird in Evaluationen voraussichtlich am meisten zurückgegriffen. Die fokussierte Betrachtung von Material zur Bestätigung von A-priori-Theorien und -Hypothesen bezeichnet Massey als attributive Daten („attributinal data“). Hierunter fasst er Fragestellungen, die nicht oder nur schwer durch direkte Fragen in den Gruppendiskussionen zu beantworten sind (wie eigene Einstellungs- und Verhaltensveränderungen). Die Frageimpulse sollten in diesen Situationen weniger direktiv sein. Aus der Art der Diskussion wird dann im Anschluss geschlossen, ob beispielsweise neue Werte im Gespräch deutlich wurden, ob eine andere Sprache bzw. Haltung angenommen wurde oder ob zentrale Konzepte eigeninitiativ von der Gruppe eingebracht und bearbeitet wurden. Ein solches Vorgehen hat sicherlich seine Stärken, wenn durch Daten aus Gruppendiskussionen die Resultate von Programmen auf der Ebene der Zielgruppen erfasst werden sollen. Eine weitere Datenart bezeichnet Massey als emergente Daten („emergent data“). Hierunter werden neue, nicht antizipierte Einsichten und Erkenntnisse verstanden, die sich durch den Austausch von Meinungen, Sichtweisen und Erfahrungen aus der Diskussion ergeben. Während die attributiven Daten eher einer deduktiven Logik zugeordnet werden, bringen die emergenten Daten induktiv unbekanntes bzw. unbeachtete Perspektiven ans Tageslicht. Die dokumentarische Methode der Interpretation, welche insbesondere bei der Evaluation von Programmen in der Jugendarbeit, Schulsozialarbeit und Organisationsentwicklung häufig als Auswertungsmethode von Gruppendiskussionen eingesetzt wird, leistet ebenfalls die Gewinnung von artikulierten, attributiven und emergenten Daten (vgl. Bohnsack/Nentwig-Gesemann 2010b). Kern der dokumentarischen Interpretation ist das Trennen und spätere In-Beziehung-Setzen des wörtlichen, immanenten Sinngehalts einerseits sowie des latenten, dokumentarischen Sinngehalts andererseits. In einem ersten Schritt wird hierfür die thematische Struktur der Diskussion rekonstruiert, also das, was die Teilnehmenden in der Gruppendiskussion sagen. In einem zweiten Schritt erfolgt die Explikation des kollektiven Rahmens, innerhalb dessen das Thema durch die Gruppe behandelt wird (vgl. Bohnsack/Nentwig-Gesemann 2007). Als ein drittes Auswertungsmodell, das allerdings vorrangig in der Sozialforschung angewandt wird, soll kurz auf die diskursive Analyse bzw. Konversationsanalyse hingewiesen werden, bei welcher die Sprachhandlungen minutiös dokumentiert und ausgewertet werden (vgl. Macnaghten/Myers 2004; Puchta/Wolff 2004; Wilkinson 2000, 2011; Wolff/Puchta 2007). Es wird hierbei davon ausgegangen, dass gerade an den Details des Gesprächs (wie Pausen, Intonation etc.) die auf den Kontext bezogenen Produktionsleistungen der Sprechenden deutlich werden. Bei der Analyse interessieren daher weniger die Inhalte der Diskutanten als vielmehr die Bedeutung, die bestimmte Sprachhandlungen in der jeweiligen Diskussionssituation haben. Da in Evaluationen jedoch immer auch die

Inhalte der Diskussion zur Beantwortung von Fragestellungen genutzt werden, wird verständlich, warum die diskurs- und konversationsanalytische Auswertung im Rahmen von Evaluationsprojekten bislang weniger zur Anwendung kommt. Die dokumentarische Methode der Interpretation sowie die konversationsanalytische Auswertung sind bislang weniger als die qualitative Inhaltsanalyse (vgl. Kuckartz 2012), die ebenso aus der Grundlagenforschung stammt, an evaluative Fragestellungen und ihre Rahmenbedingungen wie engere zeitliche und finanzielle Vorgaben angepasst worden.

## 8. Fazit

Evaluationen weisen gegenüber der Grundlagenforschung sowie der angewandten Forschung Besonderheiten auf, die insbesondere in einer höheren Praxis- und Nutzerorientierung, dem mit Evaluation verbundenen Bewertungsanspruch sowie einer Dialogorientierung gegenüber den beteiligten gesellschaftlichen Akteuren bestehen. Die drei aufgeführten Charakteristika führen dazu, dass in Evaluationen in hohem Maße kommunikative und interaktive Aufgaben zu erfüllen sind. So sind beispielsweise die Informationsbedarfe und Nutzungsinteressen der Beteiligten zu erheben, Kriterien für die Bewertung des Programms zu erarbeiten und die Anliegen und Wünsche der Stakeholder in die Gestaltung des Evaluationsprozesses einzubeziehen. Gruppendiskussionen sind in besonderem Maße geeignet, diese kommunikativen und interaktiven Aufgaben innerhalb der Evaluationspraxis zu unterstützen. Es wurde gezeigt, dass Gruppendiskussionen in Evaluationen spezifische Leistungspotenziale haben. So regen sie Teilnehmende dazu an, ihre Meinungen, Positionen und Sichtweisen zu explorieren und zu explizieren sowie sensibles und episodisches Wissen mitzuteilen. Sie fördern Lernprozesse und gemeinsame Verständnisweisen sowie in der Regel auch das Engagement für das Programm und die Evaluation. Ergänzend zu den Stärken der Methode wurden auch ihre Grenzen und wichtige Erfolgsbedingungen thematisiert. Bei der Planung und Konzeption von Gruppendiskussionen müssen Evaluierende über verschiedene Gestaltungsparameter entscheiden, wie die Stellung der Gruppendiskussion im Evaluationsdesign, die Funktion der Gruppendiskussion im Evaluationszyklus, die Klärung der zu unterstützenden Evaluationsfunktion, Anzahl, Setting und Teilnehmer(innen)zusammensetzung der Diskussionen, die Gestaltung und Moderation der Diskussion und das Auswertungsverfahren. Im Artikel sind entlang dieser Dimensionen verschiedene Gestaltungsmöglichkeiten von Gruppendiskussionen in der Evaluationspraxis aufgezeigt worden. Trotz der Vorteile, welche die Methode in Evaluationsvorhaben bietet, steht die systematische Reflexion und Fundierung der Methode in der Evaluation noch am Anfang. Es erscheint vielversprechend, zukünftig intensiver zu erforschen, mit welchen Intentionen und Vorgehensweisen Gruppendiskussionen in Evaluationsprojekten zur Anwendung gelangen und welche Herausforderungen sich in der Praxis ergeben. Auf diese Weise kann das Metho-

denwissen über Gruppendiskussionsverfahren in der Programmevaluation erweitert und ein Beitrag zur Methodenentwicklung in der Evaluation geleistet werden.

## 9. Literatur

- Abma, Tineke A. (2006): „The Practice and Politics of Responsive Evaluation“. In: *American Journal of Evaluation*, 27 (1), S. 31-43.
- Allen, Bruce Jr./Boley Cruz, Tess/Leonard, Earl/Unger, Jennifer B. (2010): „Development and Validation of a Scale to Assess Attitudes and Beliefs About Menthol Cigarettes Among African American Smokers“. In: *Evaluation & the Health Professions*, 33 (4), S. 414-436.
- Bailey, Ajay (2008): „Let’s Tell You a Story: Use of Vignettes in Focus Group Discussions on HIV/AIDS Among Migrant and Mobile Men in Goa, India“. In: Liamputtong, Pranee (Hg.): *Doing Cross-Cultural Research*. Springer Netherlands, S. 253-264.
- Balch, George I./Mertens, Donna M. (1999): „Focus Group Design and Group Dynamics: Lessons from Deaf and Hard of Hearing Participants“. In: *American Journal of Evaluation*, 20 (2), S. 265-277.
- Barbour, Rosaline S./Kitzinger, Jenny (Hg.) (1999): *Book Developing Focus Group Research. Politics, Theory and Practice*. London: Sage Publications.
- Basch, Charles (1987): „Focus Group Interview: An Underutilized Research Technique for Improving Theory and Practice in Health Education“. In: *Health Education Quarterly*, 14 (4), S. 411-448.
- Bay-Cheng, Laina Y./Livingston, Jennifer A./Fava, Nicole M. (2011): „Adolescent Girls’ Assessment and Management of Sexual Risks: Insights From Focus Group Research“. In: *Youth & Society*, 43 (3), S. 1167-1193.
- Belzile, Jacqueline A./Öberg, Gunilla (2012): „Where to Begin? Grappling With How to Use Participant Interaction in Focus Group Design“. In: *Qualitative Research*, 12 (4), S. 459-472.
- Beywl, Wolfgang (1988): *Zur Weiterentwicklung der Evaluationsmethodologie: Grundlegung, Konzeption und Anwendung eines Modells der responsiven Evaluation*. Frankfurt am Main: Peter Lang.
- Beywl, Wolfgang (2006): „Evaluationsmodelle und qualitative Methoden“. In: Flick, Uwe (Hg.): *Qualitative Evaluationsforschung*. Reinbek: Rowohlt, S. 92-134.
- Beywl, Wolfgang/Kehr, Jochen/Mäder, Susanne/Niestroj, Melanie (2007): *Evaluation Schritt für Schritt*. Münster: HIBA.
- Bloor, Michael/Frankland, Jane/Thomas, Michelle/Robson, Kate (2001): *Focus Groups in Social Research*. London: Sage Publications.
- Bohnsack, Ralf (2004): „Gruppendiskussion“. In: Flick, Uwe/von Kardoff, Ernst /Steinke, Ines (Hg.): *Qualitative Forschung. Ein Handbuch*. Reinbek: Rowohlt Taschenbuch Verlag, S. 369-384.
- Bohnsack, Ralf (2006): „Qualitative Evaluation und Handlungspraxis – Grundlagen dokumentarischer Evaluationsforschung“. In: Flick, Uwe (Hg.): *Qualitative Evaluationsforschung*. Reinbek: rowohlt’s enzyklopädie, S. 135-155.
- Bohnsack, Ralf/Nentwig-Gesemann, Iris (2006): „Dokumentarische Evaluationsforschung und Gruppendiskussionsverfahren. Am Beispiel einer Evaluationsstudie zu Peer-Mediation an Schulen“. In: Bohnsack, Ralf/Przyborski, Aglaja/Schäffer, Burkhard (Hg.): *Das Gruppendiskussionsverfahren in der Forschungspraxis*. Opladen: Barbara Budrich, S. 159-182.
- Bohnsack, Ralf/Nentwig-Gesemann, Iris (Hg.) (2007): *Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bohnsack, Ralf/Nentwig-Gesemann, Iris (2010a): „Dokumentarische Evaluationsforschung und Gruppendiskussionsverfahren. Am Beispiel einer Evaluationsstudie zu Peer-Mediation an Schulen“. In: Bohnsack, Ralf/Przyborski, Aglaja/Schäffer, Burkhard (Hg.): *Das Gruppendiskussionsverfahren in der Forschungspraxis*. Opladen & Farmington Hills: Barbara Budrich, S. 267-283.
- Bohnsack, Ralf/Nentwig-Gesemann, Iris (Hg.) (2010b): *Dokumentarische Evaluationsforschung. Theoretische Grundlagen und Beispiele aus der Praxis*. Opladen & Farmington Hills: Barbara Budrich.
- Bohnsack, Ralf/Przyborski, Aglaja/Schäffer, Burkhard (Hg.) (2006): *Das Gruppendiskussionsverfahren in der Forschungspraxis*. Opladen: Verlag Barbara Budrich.



- Bohnsack, Ralf/Przyborski, Aglaja/Schäffer, Burkhard (Hg.) (2010): Das Gruppendiskussionsverfahren in der Forschungspraxis. Opladen & Farmington Hills: Barbara Budrich.
- Bortz, Jürgen/Döring, Nicola (2002): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin: Springer.
- Buttram, Joan L. (1990): „Focus Groups: A Starting Point for Needs Assessment“. In: American Journal of Evaluation, 11 (3), S. 207-212.
- Colucci, Erminia (2007): „Focus Groups Can Be Fun“: The Use of Activity-Oriented Questions in Focus Group Discussions. In: Qualitative Health Research, 17, S. 1422-1433.
- Colucci, Erminia (2008): „On the Use of Focus Groups in Cross-Cultural Research“. In: Liamputtong, Pranee (Hg.): Doing Cross-Cultural Research. Springer Netherlands, S. 233-252.
- DeGEval – Gesellschaft für Evaluation e.V. (Hg.) (2002): Standards für Evaluation. Köln.
- Ernst, Stefanie (2006): „Die Evaluation von Qualität – Möglichkeiten und Grenzen von Gruppendiskussionsverfahren“. In: Flick, Uwe (Hg.): Qualitative Evaluationsforschung. Reinbek: Rowohlt, S. 183-213.
- Flick, Uwe (1995): Qualitative Forschung. Theorie, Methoden, Anwendung in Psychologie und Sozialwissenschaften. Reinbek: Rowohlt Taschenbuch Verlag.
- Flick, Uwe (2004): Qualitative Sozialforschung. Eine Einführung. Reinbek: Rowohlt.
- Flick, Uwe (2006a): „Interviews in der qualitativen Evaluationsforschung“. In: Flick, Uwe (Hg.): Qualitative Evaluationsforschung. Reinbek: rowohlt's enzyklopädie, S. 214-232.
- Flick, Uwe (Hg.) (2006b): Qualitative Evaluationsforschung. Reinbek: rowohlt's enzyklopädie.
- Flick, Uwe/von Kardorff, Ernst/Steinke, Ines (Hg.) (2005): Qualitative Forschung. Ein Handbuch. Reinbek: Rowohlt.
- Frey, James H./Fontana, Andrea (1991): „The Group Interview in Social Research“. In: The Social Science Journal, 28, S. 175-187.
- Friedrichs, Jürgen (1990): Methoden empirischer Sozialforschung. Opladen: Westdeutscher Verlag.
- Gaiser, Ted J. (2008): „Online Focus Groups“. In: Fielding, Nigel/Lee, Raymond M./Blank, Grant (Hg.): Online Research Methods. London: Sage, S. 290-306.
- Galloway, Kristin L. (2011): „Focus Groups in the Virtual World: Implications for the Future of Evaluation“. In: New Directions for Evaluation, 131, S. 47-52.
- Gibson, Faith (2007): „Conducting Focus Groups with Children and Young People: Strategies for Success“. In: Journal of Research in Nursing, 12 (5), S. 473-483.
- Gilfores, Javier/Granado, Alonso (1995): „Using Focus Groups in Educational Research: Exploring Teachers' Perspectives on Educational Change“. In: Evaluation Review, 19 (1), S. 84-101.
- Hammersley, Martyn (2003): „Can and Should Educational Research be Educative?“. In: Oxford Review of Education, 29, S. 3-25.
- Hansen Balle, Morten/Vedung, Evert (2010): „Theory-Based Stakeholder Evaluation“. In: American Journal of Evaluation, 31 (3), S. 295-313.
- Haubrich, Karin/Holthusen, Bernd/Struhkamp, Gerlinde (2005): „Evaluation – einige Sortierungen zu einem schillernden Begriff“. In: DJI-Bulletin 72 plus.
- Hennink, Monique M. (2007): International Focus Group Research. A Handbook for the Health and Social Science. Cambridge: Cambridge University Press.
- Hoppe, Marylin J./Wells, Elizabeth A./Morrison, Diane M./Gillmore, Mary Rogers/Wilsdon, Anthony (1995): „Using Focus Groups to Discuss Sensitive Topics with Children“. In: Evaluation Review, 19 (1), S. 102-114.
- Kamberelis, George/Dimitriadis, Greg (2008): „Focus Groups: Strategic Articulations of Pedagogy, Politics and Inquiry“. In: Denzin, Norman K./Lincoln, Yvonna S. (Hg.): Collecting and Interpreting Qualitative Materials. London: Sage, S. 375-402.
- Kamberelis, George/Dimitriadis, Greg (2011): „Focus Groups: Contingent Articulations of Pedagogy, Politics and Inquiry“. In: Denzin, Norman K./Lincoln, Yvonna S. (Hg.): The SAGE Handbook of Qualitative Research. London: Sage, S. 545-562.
- Kardoff, Ernst von (2005): „Qualitative Evaluationsforschung“. In: Flick, Uwe/von Kardoff, Ernst/Steinke, Ines (Hg.) (2005): Qualitative Forschung. Ein Handbuch. Reinbek: Rowohlt Taschenbuch Verlag, S. 238-250.

- Kardoff, Ernst von (2006): „Zur gesellschaftlichen Bedeutung und Entwicklung (qualitativer) Evaluationsforschung“. In: Flick, Uwe (Hg.) (2006): *Qualitative Evaluationsforschung*. Reinbek: Rowohlt, S. 63-91.
- Kitzinger, Jenny (1993): „Understanding Aids: Researching Audience Perceptions of Acquired Immune Deficiency Syndrome“. In: Eldrige, John (Hg.): *Getting the Message*. London and New York: Routledge.
- Kitzinger, Jenny (1994): „The Methodology of Focus Groups: The Importance of Interaction Between Research Participants“. In: *Sociology of Health & Illness*, 16 (1), S. 103-121.
- Kitzinger, Jenny (2005): „Focus Group Research: Using Group Dynamics to Explore Perceptions, Experiences and Understandings“. In: Holloway, Immy (Hg.): *Qualitative Research in Health Care*, S. 56-70.
- Kroll, Thilo/Barbour, Rosaline/Harris, Jennifer (2007): „Using Focus Groups in Disability Research“. In: *Qualitative Health Research*, 17 (5), S. 690-698.
- Kromrey, Helmut (1986): „Gruppendiskussion. Erfahrungen im Umgang mit einer weniger häufigen Methode empirischer Sozialwissenschaft“. In: Hoffmeyer-Zlotnik, Jürgen H. P. (Hg.): *Qualitative Methoden der Datenerhebung in der Arbeitsmigrantenforschung*. Mannheim: Forschung Raum u. Gesellschaft e.V., S. 109-143.
- Kromrey, Helmut (1994): *Empirische Sozialforschung*. Opladen: Leske + Budrich.
- Kromrey, Helmut (2003): „Evaluation in Wissenschaft und Gesellschaft“. In: *Zeitschrift für Evaluation*, 2, S. 93-116.
- Krueger, Richard A. (1994): *Focus Groups. A Practical Guide for Applied Research*. Thousands Oaks: Sage.
- Kuckartz, Udo (2012): *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung*. Weinheim: Beltz Juventa.
- Lambert, Sylvie D./Loiselle, Carmen G. (2008): „Combining Individual Interviews and Focus Groups to Enhance Data Richness“. In: *Journal of Advanced Nursing*, 62 (2), S. 228-237.
- Lamnek, Siegfried (1995): *Qualitative Sozialforschung. Band 2: Methoden und Techniken*. Weinheim: Beltz.
- Lamnek, Siegfried (1998): *Gruppendiskussion. Theorie und Praxis*. Weinheim: Beltz.
- Lamnek, Siegfried (2005): *Gruppendiskussion. Theorie und Praxis*. Weinheim: Beltz.
- Liamputtong, Pranee (2011): *Focus Group Methodology: Principles and Practice*. London: Sage.
- Linhorst, Donald M. (2002): „A Review of the Use and Potential of Focus Groups in Social Work Research“. In: *Qualitative Social Work*, 1, S. 208-228.
- Loos, Peter/Schäffer, Burkhard (2001): *Das Gruppendiskussionsverfahren*. Opladen: Leske + Budrich.
- Lüders, Christian (2006): „Qualitative Evaluationsforschung – Was heißt hier Forschung?“. In: Flick, Uwe (Hg.): *Qualitative Evaluationsforschung*. Reinbek: Rowohlt, S. 33-62.
- Macnaghten, Phil/Myers, Greg (2004): „Focus Groups“. In: Seale, Clive/Gobo, Giampetro/Gubrium, Jaber F./Silverman, David (Hg.): *Qualitative Research Practice*. London: Sage, S. 65-79.
- Mäder, Susanne (2005): „Fokusgruppe – Ein Instrument zur Planung und Evaluation von Stiftungsprojekten“. In: *Stiftung & Sponsoring*, S. 3.
- Mäder, Susanne/Schmidt, Stefan (2005): „Anforderungen an ein Joint-Venture in der öffentlichen Verwaltung – Empirische Befunde zu weiteren Erfolgsfaktoren der Arbeitsgemeinschaften nach § 44b SGB II“. In: *NDV-Nachrichten*, Heft 6.
- Madriz, Esther I. (1998): „Using Focus Groups with Lower Socioeconomic Status“. In: *Qualitative Inquiry*, 4 (1), S. 114-128.
- Madriz, Esther (2003): „Focus Groups in Feminist Research“. In: Denzin, Norman K./Lincoln, Yvonna S. (Hg.): *Collecting and Interpreting Qualitative Materials*. London: Sage, S. 363-388.
- Makhoul, Jihad/Nakkash, Rima (2007): „Understanding Youth: Using Qualitative Methods to Verify Quantitative Community Indicators“. In: *Health Promotion Practice*, 10, (1), S. 128-135.
- Mangold, Werner (1960): *Gegenstand und Methode des Gruppendiskussionsverfahrens*. Aus der Arbeit des Instituts für Sozialforschung. Frankfurt am Main: Europäische Verlagsanstalt.
- Mangold, Werner (1973): „Gruppendiskussion“. In: König, Rene (Hg.): *Grundlegende Methoden und Techniken der empirischen Sozialforschung*. Erster Teil. Stuttgart: Ferdinand Enke, S. 228- 259.
- Massey, Oliver T. (2011): „A Proposed Model for the Analysis and Interpretation of Focus Groups in Evaluation Research“. In: *Evaluation and Program Planning*, 34, S. 21-28.

- McClelland, Gabrielle Tracy/Newell, Robert (2008): „A Qualitative Study of the Experiences of Mothers Involved in Street-based Prostitution and Problematic Substance Use“. In: *Journal of Research in Nursing*, 13 (5), S. 437-447.
- Mensching, Anja (2006): „Zwischen Überforderung und Banalisierung – zu den Schwierigkeiten der Vermittlungsarbeit im Rahmen qualitativer Evaluationsforschung“. In: Flick, Uwe (Hg.): *Qualitative Evaluationsforschung*. Reinbek: Rowohlt, S. 339-362.
- Merton, Robert King (1987): „The Focused Interview and Focus Group: Continuities and Discontinuities“. In: *American Journal of Sociology*, 51, S. 541-557.
- Merton, Robert King/Kendall, Patricia L. (1979): „Das fokussierte Interview“. In: Hopf, Christel/Weingarten, Elmar (Hg.) (1979): *Qualitative Sozialforschung*. Stuttgart: Klett-Cotta, S. 171-204.
- Merton, Robert King/Lowenthal, Marjorie Fiske/Kendall, Patricia L. (1956, 1994): *The Focused Interview*. New York: The free Press.
- Morgan, David L. (1988): *Focus Groups as Qualitative Research*. Newbury Park: Sage.
- Morgan, David L. (1993): *Successful Focus Groups: Advancing in the State of the Art*. Newbury Park: Sage.
- Morgan, David L. (1997): *Focus Groups as Qualitative Research*. Newbury Park: Sage.
- Morgan, David L. (1998): *The Focus Group Guidebook*. London: Sage.
- Morgan, David L. (2000): „Focus Group Interviewing“. In: Gubrium, Jaber/Holstein, James A. (Hg.): *The Handbook of Interview Research*. London: Sage, S. 141-160.
- Morgan, David L. (2012): „Focus Groups and Social Interaction“. In: Gubrium, Jaber F./Holstein, James A./Marvasti, Amir/McKinney, Karyn (Hg.): *Handbook of Interview Research*. Thousand Oaks, CA: Sage, S. 161-175.
- Morgan, David L./Krueger, Richard (1998): *The Focus Group Kit*. Thousand Oaks: Sage.
- Morley, David (1996): „Medienrepublika aus der Sicht der Cultural Studies“. In: Hasenbrink, Uwe/Krotz, Friedrich (Hg.): *Die Zuschauer als Fernsehregisseure – Zum Verständnis individueller Nutzungs- und Rezeptionsmuster*. Baden-Baden, Hamburg: Nomos, S. 37-51.
- Moro, Giuseppe/Cassibba, Rosalinda/Constantini, Alessandro (2007): „Focus Groups as an Instrument to Define Evaluation Criteria“. In: *Evaluation*, 13 (3), S. 340-357.
- Myers, Greg/Macnaghten, Phil (1999): „Can Focus Groups be Analysed as Talk?“. In: Barbour, Rosaline S./Kitzinger, Jenny (Hg.): *Developing Focus Group Research*. London: Sage, S. 173-185.
- Nentwig-Gesemann, Iris (2006): „Dokumentarische Evaluationsforschung“. In: Flick, Uwe (Hg.): *Qualitative Evaluationsforschung*. Reinbek: rowohlt's enzyklopädie, S. 159-182.
- Nentwig-Gesemann, Iris (2010): „Dokumentarische Evaluationsforschung, rekonstruktive Qualitätsforschung und Perspektiven für die Qualitätsentwicklung“. In: Bohnsack, Ralf/Nentwig-Gesemann, Iris (Hg.): *Dokumentarische Evaluationsforschung. Theoretische Grundlagen und Beispiele aus der Praxis*. Opladen & Farmington Hills: Barbara Budrich, S. 63-78.
- Nentwig-Gesemann, Iris/Bohnsack, Ralf (2005): „Peer-Mediation in der Schule. Eine qualitative Evaluationsstudie zu einem Mediationsprojekt am Beispiel einer Berliner Oberschule“. In: *Deutsche Kinder- und Jugendstiftung* (Hg.): *Jung. Talentierte. Chancenreich?* Opladen: Verlag Barbara Budrich, S. 143-176.
- Nießen, Manfred (1977): *Gruppendiskussion. Interpretative Methodologie: Methodenbegründung, Anwendung*. München: Wilhelm Fink Verlag.
- O'Brien, Kerth (1993): „Using Focus Groups to Develop Health Surveys: An Example from Research on Social Relationships and Aids-Preventive Behaviour“. In: *Health Education Quarterly*, 20 (3), S. 361-372.
- Orr, Shannon K. (2010): „Exploring Stakeholder Values and Interests in Evaluation“. In: *American Journal of Evaluation*, 31 (4), S. 557-569.
- Patton, Michael Quinn (1990): *Qualitative Research & Evaluation Methods*. Thousand Oaks: Sage.
- Patton, Michael Quinn (1998): „Die Entdeckung des Prozessnutzens“. In: Heiner, Maja (Hg.): *Experimentierende Evaluation. Ansätze zur Entwicklung lernender Organisationen*. Weinheim: Juventa.
- Patton, Michael Quinn (2002): *Qualitative Research & Evaluation Methods*. Thousand Oaks u.a.: Sage.
- Peek, Lori/Fothergill, Alica (2009): „Using Focus Groups: Lessons From Studying Daycare Centers, 9/11, and Hurricane Katrina“. In: *Qualitative Research*, 9 (3), S. 31-59.

- Perreault, Michel/Renaud, Jeff/Bourassa, Francine/Beauchesne, Louise/Mpiana, André/Bernier, Sylvain/Milton, Diana (2010): „Implementation of a Panel of Service Users for the Evaluation of Mental Health Outpatient Services“. In: *Evaluation & the Health Professions*, 33 (4), S. 480-496.
- Philo, Greg/Seeker, Jenny/Platt, Steve/Henderson, Lesley/McGlaughlin, Greg/Burnside, Jocelyn (1994): „The Impact of Mass Media on Public Images of Mental Illness: Media Content and Audience Belief“. In: *Health Education Journal*, 53, S. 271-281.
- Pollock, Friedrich (1955): *Gruppenexperiment. Ein Studienbericht*. Frankfurt: Europäische Verlags-Anstalt.
- Puchta, Claudia/Wolff, Stephan (2004): „Diskursanalysen institutioneller Gespräche: Das Beispiel von ‚Focus Groups‘“. In: Keller, Reiner/Hirsland, Andreas/Schneider, Werner/Viehöver, Willy (Hg.): *Handbuch Sozialwissenschaftliche Diskursanalyse*. Opladen: Leske + Budrich, S. 439-456.
- Race, Kathryn/Hotch, Deborah/Packer, Todd (1994): „Rehabilitation Program Evaluation: Use of Focus Groups to Empower Clients“. In: *Evaluation Review*, 18 (6), S. 730-740.
- Richter, Judith M./Bottenberg, Donna J./Roberto, Karen A. (1991): „Focus Groups: Implications for Program Evaluation of Mental Health Services“. In: *The Journal of Mental Health Administration*, 18, S. 148-153.
- Robinson, Jude (2009): „Laughter and Forgetting: Using Focus Group to Discuss Smoking and Motherhood in Low-Income Areas in the UK“. In: *International Journal of Qualitative Studies in Education*, 22 (3), S. 236-278.
- Rossi, Peter Henry/Freeman, Howard/Lipsey, Mark W. (2003): *Evaluation – A Systematic Approach*. Thousand Oaks: Sage.
- Schilder, Klaas/Tomov, Toma/Mladenova, Maya/Mayeya, John/Jenkins, Rachel/Gulbinat, Walter/Manderscheid, Ron/Baingana, Florence/Whiteford, Harvey/Khandeval, Sudhir/Minoletti, Alberto/Mubbashar, Malik/Murthy, R. Srinivasa/Deva, M. Parameshvara/Baba, Aliko/Townsend, Clare/Sakuta, Tetsuya (2004): „The Appropriateness and Use of Focus Group Methodology Across International Mental Health Communities“. In: *International Review of Psychiatry*, 16, S. 24-30.
- Schwarz, Miriam/Landis, Suzanne E./Rowe, John E./Janes, Cynthia L./Pullmann, Naomi (2000): „Using Focus Groups to Assess Primary Care Patients’ Satisfaction“. In: *Evaluation & the Health Professions*, 23 (1), S. 58-71.
- Silverman, David (2004): *Qualitative Research: Theory, Method and Practice*. London: Sage.
- Smithson, Janet (2008): „Focus Groups“. In: Alasuurtari, Pertti/Bickman, Leonard/Brannen, Julia (Hg.): *The Sage Handbook of Social Research Methods*. London: Sage, S. 357-370.
- Stake, Robert E. (2004): *Standards-Based & Responsive Evaluation*. Thousand Oaks: Sage.
- Stewart, David W./Shamdasani, Prem N. (1990): *Focus Groups: Theory and Practice*. Newbury Park: Sage.
- Stockmann, Reinhard (2000): *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. Opladen: Leske + Budrich.
- Streblov, Claudia (2005): *Schulsozialarbeit und Lebenswelten Jugendlicher*. Opladen: Barbara Budrich.
- Wesnierski, Hans-Jürgen von (2003): „Rekonstruktive Sozialpädagogik im intermediären Feld eines Wissenschaft-Praxis-Diskurses. Das Beispiel Praxisforschung“. In: Schweppe, Cornelia (Hg.): *Qualitative Forschung in der Sozialpädagogik*. Opladen: Leske + Budrich, S. 69-90.
- Wibeck, Victoria/Dahlgren, Madeleine Abrandt/Öberg, Gunilla (2007): „Learning in Focus Groups: an Analytical Dimension for Enhancing Focus Group Research“. In: *Qualitative Research*, 7 (2), S. 249-267.
- Wiggins, Geoffrey Slater (2004): „The Analysis of Focus Groups in Published Research Articles“. In: *The Canadian Journal of Program Evaluation*, 19, S. 143-164.
- Wilkinson, Sue (1998a): „Focus Group Methodology: A Review“. In: *International Journal of Social Research Methodology*, 1, S. 181-203.
- Wilkinson, Sue (1998b): „Focus Groups in Feminist Research: Power, Interaction and the Co-construction of Meaning“. In: *Women’s Studies International Forum*, 21, S. 111-125.
- Wilkinson, Sue (1999a): „Focus Groups. A Feminist Method“. In: *Psychology of Women Quarterly*, 23, S. 221-224.
- Wilkinson, Sue (1999b): „How Useful are Focus Groups in Feminist Research?“. In: Barbour, Rosaline S./Kitzinger, Jenny (Hg.): *Developing Focus Group Research. Politics, Theory and Practice*. London: Sage, S. 64-75.

- Wilkinson, Sue (2000): „Women with Breast Cancer Talking Causes: Comparing Content, Biographical and Discursive Analyses“. In: *Feminism & Psychology*, 10, S. 431-460.
- Wilkinson, Sue (2004): „Focus Group Research“. In: Silverman, David (Hg.): *Qualitative Research: Theory, Method and Practice*. London: Sage, S. 177-183.
- Wilkinson, Sue (2011): „Analysing Focus Group Data“. In: Silverman, David (Hg.): *Qualitative Research: Theory, Method and Practice*. London: Sage, S. 168-184.
- Wilkinson, Clare E./Rees, Charlotte E./Knight, Lynn V. (2007): „From the Heart of My Bottom“: Negotiating Humor in Focus Group Discussions. In: *Qualitative Health Research*, 17 (3), S. 411-422.
- Willis, Karen/Green, Julie/Daly, Jeanne/Williamson, Lara/Bandyopadhyay, Mridula (2009): „Perils and Possibilities: Achieving Best Evidence From Focus Groups in Public Health Research“. In: *Australian and New Zealand Journal of Public Health*, 32 (2), S. 131-136.
- Wolff, Stephan/Puchta, Claudia (2007): *Realitäten zur Ansicht*. Stuttgart: Lucius & Lucius.
- Wottawa, Heinrich/Thierau, Heike (2003): *Lehrbuch Evaluation*. Bern: Huber.
- Wyatt, Tami H./Krauskopf, Patricia B./Davidson, Rachel (2008): „Using Focus Groups for Program Planning and Evaluation“. In: *The Journal of School Nursing*, 24 (2), S. 71-82.