

Die Ratingkonferenz

Eine Kombination von Kurzfragebogen und Gruppeninterview

Hans Keller,¹ Elke Heinemann,² Margret Kruse³

1. Einleitung

Eine Ratingkonferenz kombiniert in der Erhebungsphase quantitative und qualitative Methoden in der Form von Kurzfragebogen und direkt anschließendem Gruppeninterview. Nach der Beschreibung des Vorgehens werden die Vorteile in der Erhebungs- und Auswertungsphase dargestellt und die Grenzen aufgezeigt. Bei angemessenem Einsatz kann mit einer Ratingkonferenz mit begrenztem Aufwand ein differenziertes Bild des Evaluationsgegenstands erreicht werden. Die Methode ist im Bildungsbereich verbreitet, daher wird sie hier auch kritisch beleuchtet.

2. Die Methode

2.1 Entstehung und Entwicklung

Die Ratingkonferenz ist eine Befragung von Gruppen, die mit einem einfachen, strukturierten Fragebogen startet, der gleich nach dem individuellen Ausfüllen eingesammelt und ausgewertet wird. Anschließend werden die Ergebnisse visualisiert und mit den Befragten diskutiert. Die Ratingkonferenz ist eine Kombination von standardisierter Kurzbefragung und strukturiertem Gruppeninterview.

Die einzelnen Teile dieser Methode sind lange bekannt, weniger üblich ist diese Mischung von quantitativen und qualitativen Elementen in kompakter Form. Sie wurde anlässlich einer Evaluation des Schulklimas einer Aarauer Schule im April 1999 spontan entwickelt, mit dem Einstieg in eine Gruppendiskussion durch zwei skalierte Fragen, die mit Klebepunkten zu beantworten waren.

¹ Institut für Gymnasial- und Berufspädagogik, Universität Zürich

² Freiberufliche Kommunikationswissenschaftlerin, Bremen

³ Univation Institut für Evaluation Dr. Beywl & Associates GmbH, Köln, Hamburg

Ausgehend von dieser Erfahrung wurde die oben kurz beschriebene Methode⁴ im Rahmen der Entwicklung des Q2E-Systems der Schulqualitätsentwicklung und -evaluation (www.q2e.ch) verfeinert. Norbert Landwehr (2003c: 60f.) prägte den Begriff der „Ratingkonferenz“ (Fragen mit Rating-Skalen und anschließender Diskussion).

Die Methode wird seit den ersten externen Schulevaluationen nach Q2E in diesem Anwendungsfeld breit eingesetzt, insbesondere vom Institut für externe Schulevaluationen der Sekundarstufe II (ifes Zürich). Weiter zeigte sich in der Q2E-Entwicklung, dass die Ratingkonferenz auch sehr gut geeignet ist für das Einholen von Feedback durch Lehrpersonen und Dozierende in ihren Kursen (vgl. Landwehr 2005; Keller 2005; Beywl/Bestvater/Friedrich 2011).⁵

Darüber hinaus hat sich diese Methode auch in anderen Evaluationen im Bildungsbereich bewährt.

2.2 Beschreibung der Methode

Die Methode bietet sich an, wenn viele Personen zu befragen sind. Die Gruppengröße bei einer Ratingkonferenz beträgt in der Regel acht bis zwölf Personen. Entsprechend der Grundgesamtheit werden mehrere Gruppen zusammengestellt. Eine Ratingkonferenz dauert zwischen 60 und 90 Minuten.

Der Fragebogen umfasst sechs bis zehn zentrale, sorgfältig formulierte Fragen mit skalierten Antwortvorgaben und i.d.R. eine oder zwei offene Fragen. Die Antworten auf die offenen Fragen werden mittels Moderationskarten aufgenommen oder protokolliert, während die eingesammelten Fragebogen ausgewertet (ausgezählt) werden.

Die Ergebnisse werden anschließend offen visualisiert und die Teilnehmenden werden um Ergänzungen und Erläuterungen gebeten. Die Kommentare werden ebenfalls offen protokolliert (Flipchart oder PC/Beamer), so können die Befragten, wenn erforderlich, intervenieren.

Das Ziel der Gespräche ist die Verdeutlichung der Antworten und Klärung von Differenzen und Übereinstimmungen, d.h. die kommunikative Validierung und Anreicherung der quantitativen Antworten, nicht eine Einigung der Befragten.

Die Auswertung der Erhebung erfolgt primär qualitativ, die quantitativen Daten dienen lediglich zur Unterstützung der qualitativen Aussagen, nicht für eigenständige Ergebnisse. Die Verteilung der quantitativen Antworten zu jeder Frage wird vor allem als Argumentationshintergrund erstellt. Weitere statistische Analysen sind bei den begrenzten Daten wenig sinnvoll. Hingegen lohnt sich die Gegenüber-

4 Eine ausführliche Ablaufbeschreibung einer Ratingkonferenz sowie ein Fragebogenbeispiel aus einer Evaluation befinden sich im Anhang.

5 Im Unterschied zu einer externen oder Fremdevaluation befragt hier meist die betreffende Person selbst ihre Klasse(n), setzt die Methode also im Rahmen einer Selbstevaluation ein. Bei dieser Anwendung steht der Moderationscharakter im Vordergrund. Die Lehrperson muss darauf achten, nicht schon in eine Diskussion einzutreten, sondern zuerst die Rückmeldungen aufzunehmen und zu klären. Anonymität besteht hier nicht. Ein Protokoll der Gespräche entfällt, dafür werden allfällige Umsetzungen der Feedbacks besprochen und festgehalten.

stellung der Antworten verschiedener Gruppen, insbesondere wenn sich diese in ihrer Zusammensetzung unterscheiden. So haben sich z.B. Gruppen nur mit Lernenden bzw. Lehrpersonen, idealerweise je mindestens zwei Gruppen (Triangulation), bewährt.

Die Protokolle und die Ergebnisse der Moderation der offenen Frage(n) werden im Sinne der zusammenfassenden Inhaltsanalyse (vgl. Mayring 2010: 63ff.) resümiert und zentrale Aussagen möglichst durch Zitate aus den Interviews oder quantitativen Daten aus dem Fragebogen belegt.

Die Ratingkonferenz kombiniert quantitative und qualitative Erhebungsformen, so dass die Schwächen dieser Erhebungsverfahren nicht vermieden, aber ausgeglichen werden können. Durch den Fragebogen ist im Gegensatz zu offeneren Verfahren eine gewisse Vergleichbarkeit und Standardisierung gegeben und alle Teilnehmenden werden zu Statements angehalten.

Die Interviewenden moderieren hier wie in Leitfaden-Interviews und achten darauf, alle am Gespräch zu beteiligen. Die Befragten können sich nicht nur zu den gestellten Fragen äußern, sondern auch zu von den Interviewenden nicht thematisierte, für sie aber relevante Aspekte in das Gespräch einbringen. Dadurch wird das Erhebungsverfahren offener und es dient zudem „zur Erfassung der in quantitativen Befragungen kaum zu ermittelnden subjektiven Deutungsmuster (...)“ (Ernst 2006: 195).

Bei der Anwendung der Ratingkonferenzen werden die Standards der Gesellschaft für Evaluation (vgl. DeGEval 2008) beachtet: Die einzelnen Kriterien der Anforderungen an eine Evaluation in den Bereichen Nützlichkeit, Durchführbarkeit, Genauigkeit und Fairness sind handlungsleitend, wobei bei der Auswahl der Ratingkonferenz als Erhebungsmethode besonderer Wert auf den Aspekt der Nützlichkeit gelegt wird, ohne die anderen zu vernachlässigen.

3. Voraussetzungen

3.1 Einsatzbereich

Die Methode eignet sich für Erhebungen, in deren Rahmen einerseits mehrere Gruppen unter vergleichbaren Voraussetzungen zu einem festgelegten Themenspektrum befragt werden sollen, andererseits jedoch die Offenheit gegenüber weiteren Aspekten aus dem Kreis der Befragten gewahrt bleiben soll. Die Durchführung von Ratingkonferenzen ist angebracht, wenn es nicht auf die Erfassung von Einzelmeinungen ankommt, die Ressourcen gering sind oder der Aufwand für Einzelinterviews der Fragestellung nicht angemessen ist.

Im Rahmen von internen und externen Evaluationen bieten Ratingkonferenzen die Möglichkeit, Meinungen, Erfahrungen und Wissen der Befragten zu einem bestimmten durch den Fragebogen vorgegebenen Themenspektrum zu ermitteln. Dieses Spektrum kann von den Befragten während der Ratingkonferenz erweitert werden. Gezielte Nachfragen der Moderierenden können die Teilnehmenden zudem

zu detaillierten Erläuterungen anregen, Differenzen werden so nicht nur als Zahl sichtbar, sondern verbal durch die Teilnehmenden präzisiert. Erwartete und auch unerwartete Aspekte werden so thematisiert und vorab nicht erkannte Gesichtspunkte und Entwicklungsfelder können erörtert werden (vgl. Kelle 2006: 131ff.). Dies könnte auch durch Leitfaden-Interviews erreicht werden, aber nur mit kleineren Teilnehmendenzahlen.

Da der Erhebungsaufwand gegenüber anderen Interviewformen begrenzt ist, können mehrere Gruppen zum gleichen Thema befragt werden und so kann ein mehrperspektivischer Blick auf das Thema geworfen werden. Durch die Gegenüberstellung der Ergebnisse der verschiedenen Gruppen und der quantitativ und qualitativ erhobenen Daten gewinnen die Ergebnisse an Validität (vgl. Kromrey 2002: 524f.).

3.2 Fragestellung

Ausgehend von der Forschungsfrage bzw. Evaluationsfragestellung werden für den Fragebogen einige wenige zentrale Fragen ausgewählt, die den Themenbereich ausleuchten. Dies ist auch in anderen Gruppeninterviews möglich, aber die Ratingkonferenz bietet den Vorteil, dass sich die Teilnehmenden beim Ausfüllen des Fragebogens für sich, ohne Kenntnis der Gruppenmeinung, Gedanken zum Thema machen. Dabei sollte der Fragebogen der sprachlichen Kompetenz der Befragten entsprechen. Mit den Fragen muss im Gegensatz zu einer rein quantitativen Fragebogenbefragung nicht alles im Detail erfragt werden, es reicht vielmehr, die wichtigsten Aspekte aufzunehmen. Die Fragen dienen gleichzeitig als Leitfaden für das anschließende Gespräch über die Ergebnisse und strukturieren den Interviewverlauf.

Die Moderierenden können vorab im Evaluationsteam Nachfragen vereinbaren, die dann in allen zu befragenden Gruppen gestellt werden; des Weiteren kann spontan nachgefragt werden, wenn sich während der Ratingkonferenz neue Aspekte ergeben.

3.3 Auswahl und Zusammensetzung der Befragten

Gegenüber dem Einzelinterview bietet die Ratingkonferenz die Chance, mehrere Personen in Gruppen zum gleichen Thema zu befragen und dennoch Aussagen und Interpretationen Einzelner einzubeziehen. Die ideale Gruppengröße liegt, wie beschrieben, bei acht bis zwölf Teilnehmenden, da so für die Einzelnen die Möglichkeit der Partizipation eher gegeben ist.

In kleineren Organisationseinheiten ermöglicht die Methode der Ratingkonferenz eine Vollbefragung, da mit relativ geringerem Aufwand viele Befragte erfasst werden können. In größeren Organisationseinheiten werden Stichproben gezogen, so dass Betroffene aus allen relevanten Bereichen beteiligt werden. Die Befragten können als Experten und Expertinnen für die zu untersuchenden Themen angesehen werden, da sie diesbezüglich einen Wissensvorsprung besitzen, der aus ihrer Positi-

on in dem spezifischen Kontext resultiert (vgl. Gläser/Laudel 2010: 11ff.; Bohnsack/Marotzki/Meuser 2011: 57f.).

Eine wesentliche Voraussetzung, nicht nur bei Ratingkonferenzen, sondern generell sowohl bei quantitativen als auch bei qualitativen Befragungsverfahren, ist die kommunikative Kompetenz der Beteiligten. Diese müssen die Begriffe, Grammatik und Semantik nicht nur kennen, sondern ihnen auch die gleiche Bedeutung zumessen; das ist unabdingbar für einen gelingenden Kommunikationsprozess (vgl. Friedrichs 1981: 207; Kromrey 2002: 356; Atteslander 2003: 126f.; Helferich 2011: 21f.). Die Methode erlaubt es auch, Teilnehmende mit geringerer sprachlicher Kompetenz zu befragen, indem einerseits der Fragebogen ihren sprachlichen Möglichkeiten angepasst wird und zudem im Gespräch Raum für Erklärungen ist.

3.4 Gesprächsleitung

An der Durchführung der Ratingkonferenz, der späteren Auswertung sowie der Berichterstellung sind in der Regel zwei Personen beteiligt.⁶ Eine übernimmt die Moderation und die andere die Protokollführung. Die Auswertung erfolgt gemeinsam. Bei größeren Evaluationen wird ein Team aus mindestens vier Personen gebildet, damit parallele Befragungen möglich sind. Dies hat sich besonders bei externen Schulevaluationen (Q2E, ifes und mehrere Schweizer kantonale Evaluationsstellen) bewährt.

Der interviewleitenden Person obliegt die Aufgabe, durch ihre Moderation eine möglichst natürliche sowie vertrauensvolle Atmosphäre zu schaffen und die Teilnehmenden auch miteinander ins Gespräch zu bringen. Dabei ist zu berücksichtigen, ob die Teilnehmenden sich bereits kennen und inwieweit dies die Offenheit der Gespräche beeinflusst.

Die moderierende Person sollte den Interviewverlauf flexibel gestalten, also nicht strikt nach Fragebogenablauf vorgehen und eine Einflussnahme durch eigene inhaltliche verbale oder nonverbale Äußerungen auf den Gesprächsverlauf weitgehend vermeiden. Weiter ist zu versuchen, möglichst alle Teilnehmenden einzubeziehen und die Dominanz einzelner Gruppenmitglieder einzuschränken (vgl. Lamnek 2010: 399ff.). Werden die Ratingkonferenzen von verschiedenen Personen durchgeführt, stimmen diese ihr Vorgehen möglichst präzise ab, um die Vergleichbarkeit zu gewährleisten. Letztendlich verlaufen die Interviewsituationen und somit die Datengewinnung stets unterschiedlich und die Interviewenden werden z.B. allein durch ihr Auftreten und Handeln immer einen gewissen Einfluss auf das Geschehen ausüben (vgl. Atteslander 2003: 123ff., 152f.; Kromrey 2002: 358). „Auch durch noch so intensive Interviewschulungen und durch Standardisierungsbemühungen lässt sich damit kein Messinstrument entwickeln, das von Situationseinflüssen unbeeinflusste, im methodischen Sinn ‚zuverlässige‘ Daten produziert“ (Kromrey 2002: 394).

⁶ Landwehr sieht einen Vorteil darin, wenn nur eine Person moderiert und protokolliert: Das Protokoll werde so mehr beachtet (vgl. Landwehr 2005).

4. Stärken⁷

4.1 Kommunikative Validierung

Da die Ergebnisse mit den Befragten besprochen werden, erübrigen sich spekulative Interpretationen der Ergebnisse insbesondere zu den Rating-Items durch das Evaluationsteam. Oft liefern die Befragten Begründungen und zusätzliche Informationen zu ihren Antworten. Die Methode profitiert gegenüber rein quantitativen Erhebungen von der Chance des Nachfragens. Die Befragten kommen im Gespräch selbst zu Wort und „(...) sind zunächst die Experten für ihre eigenen Bedeutungsgehalte“ (Mayring 1990: 45).

Wie jede schriftliche Befragung stößt auch der Fragebogen in der Ratingkonferenz auf Grenzen der sprachlichen Fähigkeiten und des sprachlichen Verständnisses der Befragten (siehe 5.1). Gerade dann zeigt sich die Stärke der Methode: Schon beim Ausfüllen des Fragebogens können die Evaluierenden das Verhalten der Teilnehmenden beobachten und nötigenfalls vorweg die Fragen und Begriffe erklären oder bei Bedarf alle durch den Fragebogen führen (dabei ist besonders sorgfältig darauf zu achten, die Antworten nicht zu beeinflussen). Außerdem zeigen sich im Gespräch mögliche Schwierigkeiten der Beantwortung sofort und können aufgefangen werden.

Ein weiterer Vorteil ist, wie bereits erwähnt, dass Unklarheiten und Missverständnisse in der Frageformulierung auch im anschließenden Gespräch geklärt werden können. Oft zeigt sich, dass versehentlich falsch angekreuzt wurde. Einzelne Befragte korrigieren manchmal sogar ihre Einschätzung aufgrund der Klärung. „Durch qualitative Daten lassen sich also Validitätsbedrohungen und Methodenprobleme standardisierter Befragungen aufdecken, und auf diese Weise kann dann oftmals eine realistischere Einschätzung des Informationsgehalts bestimmter empirischer Daten in Evaluationsstudien gewonnen werden“ (Kelle 2006: 133).

4.2 Aufwand – Ertrag

In Evaluationen ist die Durchführung von Ratingkonferenzen angebracht, wenn Einzelinterviews aus unterschiedlichen Gründen (Zeit, Kosten) nicht zu realisieren sind und eine rein quantitative Befragung zu wenig reichhaltige Informationen verspricht. Durch eine geeignete Auswahl der Befragten (siehe 3.3) kann mit begrenztem Aufwand ein gutes Bild der unterschiedlichen Sichtweisen zur Beschreibung der Situation erstellt werden.

Ein positiver Nebeneffekt der Methode liegt im immer wieder zu beobachtenden Entwicklungsaspekt: Die Teilnehmenden tauschen ihre Erfahrungen und ihr Wissen aus und entwickeln Ideen für die eigene Praxis. So können sich aus den

7 Zu Stärken und Grenzen der Ratingkonferenz in der Lehrevaluation siehe Landwehr (2005).

Ratingkonferenzen Sekundäreffekte für die Organisations- und Personalentwicklung ergeben.

4.3 Methoden-Mix

Die Kombination von quantitativen und qualitativen Elementen hat sich bewährt, da „jedes Verfahren spezifische Stärken und Schwächen aufweist und Sachverhalte (...) mit unterschiedlicher Breite und Tiefe erfasst“ (Kromrey 2002: 524).

Mit dem Fragebogen – und allenfalls mit den Moderationskarten zur offenen Frage – kommt zu Beginn jede/-r Beteiligte „zu Wort“. Die gesammelten Ergebnisse zeigen ein Meinungsbild und erste Trends für die ganze Gruppe bevor erste Äußerungen das Gespräch in bestimmte Richtungen führen. Alle haben zudem die Gelegenheit, im anschließenden Gespräch ihre jeweilige Ansicht zu verdeutlichen.

Das Gespräch ist durch die Fragen des Fragebogens vorstrukturiert und jede/-r Befragte hat sich beim Ankreuzen schon eigene Gedanken gemacht. Dadurch erhält das Gespräch eine transparente Struktur, ohne dass die Fragen sklavisch abgearbeitet werden müssen. Die zentralen Diskussionspunkte werden durch jene Fragen markiert, bei denen die Antworten streuen, die Teilnehmenden von sich aus Erläuterungen geben oder bei denen alle eine gewisse Unzufriedenheit ausdrücken.

Als Ergebnis hat man am Schluss eine quantitative Übersicht sowie Erläuterungen und Ergänzungen zu den Antworten, die insgesamt die Zusammenfassung und Interpretation erleichtern.

5. Grenzen der Methode

5.1 Sprachliche/soziolinguistische Aspekte

Wie bei allen Interviewformen gilt es, die sprachlichen Unzulänglichkeiten zu beachten. „Die Methode setzt gute verbale Fähigkeiten der Befragten und ein weitgehend ähnliches Sprachspiel (Wittgenstein) bei Forscher, Interviewer und Befragten voraus“ (Friedrichs 1981: 207).

Gerade wenn man glaubt, gute Feldkenntnisse zu haben, ist Vorsicht bei der Interpretation angebracht. Die sprachlichen Äußerungen sind an den konkreten Kontext gebunden (vgl. Lamnek 2010: 20f.; Helfferich 2011: 22). „(...) [N]icht gesichert ist, dass verschiedene Personen unter der gleichen Frage (dem gleichen Fragewortlaut) das gleiche verstehen“ (Kromrey 2002: 396). Dieser Effekt zeigt sich immer wieder in den Ratingkonferenzen bei der Besprechung der Antworten, wo diese Differenzen dann geklärt werden können.

5.2 Begrenzte Repräsentativität

Die Ratingkonferenz wird in der Regel mit einer Stichprobe aus der Grundgesamtheit durchgeführt. Damit stellen sich die bekannten Probleme der Auswahl und Repräsentativität.

Die quantitativen Daten allein sind nicht aussagekräftig, da im Fragebogen nur wenige Fragen gestellt werden und sie erfüllen nicht die Anforderungen für statistisch robuste Aussagen. Sie dienen nur im Zusammenhang mit den Gesprächsprotokollen als Grundlage und Beleg.

In den geeigneten Anwendungsfeldern (siehe 3.1/3.3) steht die Repräsentativität nicht an erster Stelle, sondern die Reichhaltigkeit der unterschiedlichen Perspektiven.

5.3 Gruppensituation

Während des Gesprächs über die offene(n) Frage(n), spätestens wenn die Diskussion über die Ergebnisse aus dem Fragebogen einsetzt, beginnen die Interaktionen unter den Befragten, die sich oft mehr oder weniger gut kennen. Hier kommen Einstellungen und latente Meinungen zum Vorschein (vgl. Ernst 2006: 196; Nentwig-Gesemann 2006: 160ff.; Mayring 1990: 54). Aufgrund der sozialen Beziehungen beeinflussen sich die Befragten gegenseitig (vgl. Friedrichs 1981: 215) und die Äußerungen sind gefiltert (fehlende Anonymität, soziale Erwünschtheit; vgl. Kromrey 2002: 397). Dies ist für die außenstehende Gesprächsleitung nicht immer leicht erkennbar.

Im Gruppengespräch geht es darum, diese Aspekte zu beachten und bei der Auswertung einzelne Aussagen mit entsprechender Vorsicht zu gewichten.

5.4 Protokollierung

Das während des Gesprächs erstellte und für alle Teilnehmenden sichtbare Protokoll hat den Vorteil einer ersten Verdichtung der Aussagen, birgt dafür die Gefahr in sich, die Aussagen zu verändern oder gar zu verfälschen. Das stellt hohe Anforderungen an die Protokollierenden. Die Einsehbarkeit der Protokolle während der Befragung und die Tonaufnahme⁸ sind notwendige Sicherungen der Originalaussagen.

Trotzdem können Äußerungen nur in den seltensten Fällen verbal eindeutig erfasst und wiedergegeben werden, zudem lassen sich die Voten meist nicht mehr den einzelnen Personen zuordnen.

⁸ In Absprache mit den Teilnehmenden kann das Gespräch zur Sicherheit für die Auswertung aufgezeichnet werden. Dieser Mitschnitt wird in der Regel nicht transkribiert.

5.5 Auswertung

Die Auswertung der Protokolle erfolgt nach den Methoden der qualitativen Sozialforschung, mit allen bekannten Einschränkungen: „Qualitative Forschung wird zu einem kontinuierlichem Prozess der Konstruktion von Versionen der Wirklichkeit – die Version, die jemand in einem Interview erzählt, muss nicht der Version entsprechen, die er zum Zeitpunkt des Geschehens formuliert hätte. Sie muss auch nicht der Version entsprechen, die er einem anderen Forscher mit anderer Fragestellung präsentiert hätte. Auch der Forscher, der dieses Interview auswertet und als Teil seiner Ergebnisse darstellt, produziert eine neue Version des Ganzen“ (Flick 1995: 19f.).

Von Vorteil bei den Ratingkonferenzen ist, dass sowohl die Antworten (Versionen) der verschiedenen Befragten aus den unterschiedlichen Gruppen das Gesamtbild formen, als auch, dass mindestens zwei Personen (Moderator/-in und Protokollant/-in) gemeinsam die Daten auswerten.

In der gemeinsamen Auswertung von Moderator/-in und Protokollant/-in hat es sich bewährt, die Ergebnisse zu ‚Kernaussagen‘ zusammenzufassen und diese mit detaillierteren Aussagen aus den Protokollen und den Auszählungen der Kurzbefragung zu begründen.

Wichtig ist, wie schon oben ausgeführt, die Vorsicht bei der Darstellung und Interpretation der quantitativen Daten. Sie dürfen nicht überbewertet werden und sollten eher nur auszugsweise in der Auswertung bzw. dem Bericht zitiert werden.

6. Fazit

Die Methode bewährt sich seit vielen Jahren bei den externen Evaluationen von Schulen sowie generell bei Evaluationen im Bildungsbereich, wenn Einzelinterviews oder auch freie Gruppendiskussionen aus unterschiedlichen Gründen nicht möglich oder sinnvoll sind und rein quantitative Verfahren zu wenig Erkenntnisgewinn versprechen.

Vorteile der Methode sind der relativ geringe Erhebungsaufwand, die Mischung aus quantitativen und qualitativen Daten und insbesondere die kommunikative Validierung durch die Befragten. Der Auswertungsaufwand, der wie bei allen qualitativen Erhebungen groß werden kann, lässt sich aufgrund der stärkeren Vorstrukturierung durch den Einsatz eines Kurzfragebogens und des Verzichts auf Transkription der Tonprotokolle begrenzen.

Hervorzuheben ist, dass bei entsprechend professioneller Anwendung die Mehrperspektivität und Triangulation gegeben sind und insbesondere dem Bereich Nützlichkeit der Standards für Evaluation Rechnung getragen wird.

7. Literatur

- Atteslander, Peter (2003): Methoden der empirischen Sozialforschung. Berlin: de Gruyter (10. Aufl.).
- Beywl, Wolfgang/Bestvater, Hanne/Friedrich, Verena (2011): Selbstevaluation in der Lehre. Münster u.a.: Waxmann.
- Bohnsack, Ralf/Marotzki, Winfried/Meuser, Michael (Hg.) (2011): Hauptbegriffe qualitativer Sozialforschung. Opladen: Budrich (3. Aufl.).
- DeGEval – Gesellschaft für Evaluation e.V. (Hg.) (2008): Standards der Evaluation. Mainz: Eigenverlag (4. unveränderte Auflage).
- Ernst, Stefanie (2006): Die Evaluation von Qualität – Möglichkeiten und Grenzen von Gruppendiskussionsverfahren. In: Flick, Uwe (Hg.): Qualitative Evaluationsforschung: Konzepte – Methoden – Umsetzung. Hamburg: Rowohlt, S. 183-213.
- Flick, Uwe (Hg.) (1995): Qualitative Forschung. Hamburg: Rowohlt.
- Flick, Uwe (2006): Interviews in der qualitativen Evaluationsforschung. In: Flick, Uwe (Hg.): Qualitative Evaluationsforschung. Hamburg: Rowohlt, S. 214-232.
- Friedrichs, Jürgen (1981): Methoden empirischer Sozialforschung. Opladen: Westdeutscher Verlag (9. Aufl.).
- Gläser, Jochen/Laudel, Grit (2010): Experteninterview und qualitative Inhaltsanalyse. Wiesbaden: VS Verlag (4. Aufl.).
- Helfferrich, Cornelia (2011): Die Qualität qualitativer Daten. Wiesbaden: VS Verlag (4. Aufl.).
- Kelle, Udo (2006): Qualitative Evaluationsforschung und das Kausalitätsparadigma. In: Flick, Uwe (Hg.): Qualitative Evaluationsforschung. Hamburg: Rowohlt, S. 117-134.
- Keller, Hans (2005): Aufbau und Elemente einer Feedbackkultur. Zürich: Impulse (5. Aufl.).
- Kromrey, Helmut (2002): Empirische Sozialforschung. Opladen: Budrich (10. Aufl.).
- Lamnek, Siegfried (2010): Qualitative Sozialforschung. Weinheim, Basel: Beltz (5. Aufl.).
- Landwehr, Norbert (2003c): Q2E Band 3 – Grundlagen zum Aufbau einer Feedbackkultur. Bern: hep.
- Landwehr, Norbert (2005): Lehrevaluation als Anstoss zur Unterrichtsentwicklung. In: Beiträge zur Lehrerbildung, Bd. 3/2005, S. 321-333.
- Mayring, Philipp (1990): Einführung in die qualitative Sozialforschung. München: Psychologie Verlags Union.
- Mayring, Philipp (2010): Qualitative Inhaltsanalyse. Weinheim und Basel: Beltz (11. Aufl.).
- Nentwig-Gesemann, Iris (2006): Dokumentarische Evaluationsforschung. In: Flick, Uwe (Hg.): Qualitative Evaluationsforschung. Hamburg: Rowohlt, S. 159-182.

8. Anhang

8.1 Beschreibung der Methode: Detaillierter Ablaufplan

Die Handlungsanweisungen für die jeweilige Ratingkonferenz können in einem Drehbuch für die Evaluationsteams festgehalten werden, um eine gewisse Standardisierung beim Vorgehen sicherzustellen. Für die Befragung wird wie folgt vorgegangen (Gesamtdauer 60-90 Minuten):

- Auf der Basis der Fragestellung wird ein kurzer Fragebogen (sechs bis zehn zentrale Fragen) mit skalierten Antworten erstellt. Am Ende des Fragebogens können noch ein bis zwei offene Fragen gestellt werden. Ein Beispiel aus einer Evaluation findet sich im Anhang 8.2. Es hat sich, insbesondere wenn mehrere Gruppen parallel befragt werden, bewährt, einen internen Bogen mit abgesprochenen Nachfragen zu erstellen. Der Fragebogen wird in entsprechender Anzahl kopiert.
- Die befragte Gruppe wird über das Ziel der Befragung, den Umgang mit den Daten und das Vorgehen informiert, dann werden die Fragebogen individuell ausgefüllt. Zwei Personen leiten die Befragung, eine als Gesprächsleitung, die andere unterstützend (Auszählung, Protokoll).
- Die Antworten werden gesammelt mit dem Ziel, die Ergebnisse der Gruppe rasch sichtbar zu machen. Dazu gibt es mehrere Möglichkeiten, hier drei davon:
 - a) Klebepunkte
Die Befragten übertragen ihre Antworten mit Klebepunkten auf ein vorbereitetes Plakat, das genügend große Antwortfelder für alle Punkte enthält. Dazu erhalten die Befragten je gleich viele Klebepunkte wie es Fragen auf dem Fragebogen hat.
 - b) Overhead-Folie
Die Befragten geben ihre Fragebogen ab. Diese werden auf einer Folienkopie des Fragebogens zusammengetragen. Diese Folie mit der Verteilung aller Antworten wird anschließend projiziert.
 - c) PC-Erfassung
Mit einem entsprechend vorbereiteten Programm (z.B. Tabellenkalkulation) werden die Antworten erfasst. Die Zusammenfassung (Verteilung der Antworten, nicht Mittelwerte) wird mit Beamer projiziert.
- Während die Antworten von der unterstützenden Person erfasst werden, kann der/die Gesprächsleiter(in) zwischenzeitlich die Antworten zur offenen Frage mit Moderationskarte⁹ aufnehmen, gruppieren und kurz erläutern lassen. Ergänzungen werden notiert (Flipchart).
- Liegt die Auswertung vor, so werden die Befragten um Ergänzungen und Kommentare zu den Antworten gebeten. Dabei geht es um Klärung von Diffe-

⁹ Aufgrund der schriftsprachlichen Kompetenz kann nicht in allen Gruppen mit Moderationskarten gearbeitet werden, hier bietet sich ein Gespräch zu einer offenen Frage an.

renzen und Übereinstimmungen, nicht um eine Einigung. Das Ziel ist die kommunikative Validierung und Anreicherung der quantitativen Antworten.

- Das Gespräch wird sichtbar protokolliert: Die unterstützende Person erstellt ein Protokoll. Dieses kann als öffentliches Protokoll auf einer Flipchart oder auf einem PC mit Beamer sichtbar gemacht werden. So können die Befragten direkt intervenieren, wenn ihr Beitrag nicht sinngemäß notiert worden ist, und der/die Protokollführer/-in kann auch nachfragen, ob ihre Formulierung angemessen ist. Für die weitere Verarbeitung wird gleichzeitig eine Tonaufnahme erstellt.¹⁰
- Das Gespräch wird abgeschlossen mit der Möglichkeit, nicht Besprochenes noch aufzunehmen. Die Befragten werden informiert, wie weiter mit den Antworten umgegangen wird (Anonymität, Termine, Berichterstattung).
- Die Fragebogen, das Protokoll und die Tonaufnahme werden entsprechend des Auftrags und der Ressourcen ausgewertet. In der Regel dient die Tonaufnahme nur zur Kontrolle und Ergänzung des Protokolls und für wörtliche Zitate, sie wird selten vollständig transkribiert.

8.2 Beispiel eines Fragebogens für Mitarbeiterinnen und Mitarbeiter eines Ausbildungsprojekts

Aussage					
	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	keine Angabe
Allgemein zum Projekt und zur Kooperation					
1. Das Konzept für das individualisierte und selbstgesteuerte Lernen ist eine gute Grundlage und Leitlinie für die Ausbildung.					
2. Die Unterstützung bei der Konzeptweiterentwicklung und Umsetzung ist zufriedenstellend.					
3. Vereinbarungen und Absprachen zwischen allen Beteiligten werden eingehalten.					
Neue Methoden und Instrumente					
4. Die neu entwickelten Methoden und Instrumente sind nützlich für die Ausbildung in Theorie und Praxis.					
5. Die regelmäßigen Reflexionsrunden tragen wesentlich zur Verbesserung der überfachlichen Fähigkeiten der Teilnehmer/-innen bei.					
6. Das Lernklima in der Gruppe hat sich durch die neuen Methoden und Instrumente deutlich verbessert.					

Offene Frage:

Inwieweit sind die Teilnehmerinnen und Teilnehmer besser auf die Praxis vorbereitet als frühere Jahrgänge und inwieweit beobachten Sie, dass diese selbstverantwortlicher lernen?

¹⁰ Zu Beginn der Ratingkonferenz ist die Zustimmung der Gruppe zur Aufzeichnung des Gesprächs einzuholen.