

# Methodische Herausforderungen der Wirkungsanalyse bei knappen Ressourcen

## Dokumentation der Frühjahrstagung 2015 des AK Soziale Dienstleistungen<sup>1</sup>

### Tagungsthema und Format

Wirkungsanalysen sind eine zentrale Evaluationsaufgabe, aber auch eine große methodische Herausforderung, weil die Effekte oft klein sind, unterschiedliche Ursachen haben können und häufig nicht eindeutig auf bestimmte Maßnahmen zurückzuführen sind. Die Schwierigkeiten nehmen drastisch zu, wenn die zeitlichen und materiellen Ressourcen für die Planung und Entwicklung der Instrumente sowie für die Erhebung, Datenanalyse und Ergebnisaufbereitung knapp sind. Auf Evaluationen trifft das fast immer zu. Was können Evaluatorinnen und Evaluatoren also tun, wenn es keine passende Skala für ein bestimmtes Phänomen gibt und eine aufwändige Skalenentwicklung mitunter aus Ressourcengründen nicht in Frage kommt? Oder wie soll man damit umgehen, wenn eine Randomisierung nicht machbar ist und der Datenschutz die Erhebung wichtiger Kontrollvariablen in einem quasi-experimentellen Design unmöglich macht? Wie kann man Ergebnisse aus standardisiert erhobenen Daten z. B. durch die Erhebung zusätzlicher qualitativer Daten absichern? Kann man mit qualitativen Interviews, Fokusgruppen und Gruppendiskussionen die Wirkungsfrage beantworten? Und wie lässt sich der Aufwand dabei in Grenzen halten? Geht das dann zu Lasten der methodischen Standards (oder auch der Evaluationsstandards)? Bis zu welchem Grad können oder müssen diese Standards bei Evaluationen aus dem Bereich der Sozialen Dienstleistungen eingehalten werden?

Die Frühjahrstagung des Arbeitskreises Soziale Dienstleistungen (AK) der DeGEval am 24. Mai 2015 bot Evaluatorinnen und Evaluatoren eine Gelegenheit, diese und ähnliche Fragen zu diskutieren und gemeinsam über Lösungen nachzudenken.<sup>2</sup> Rund 30 Personen aus Deutschland und Österreich nahmen im Internationalen Jahr der Evaluation an der Frühjahrstagung in Wolfsburg teil.

---

<sup>1</sup> Die Dokumentation wurde von Stefanie Reiter (Deutsches Jugendinstitut e.V.), Stefan Schmidt (schmidt evaluation), und PD Dr. Rainer Strobl (proVal, Gesellschaft für Sozialwissenschaftliche Analyse – Beratung – Evaluation) unter Mitwirkung von Mag. Florentina Astleithner (FH Campus Wien), Benjamin Froncek (FernUniversität in Hagen) und Mag. Peter Stepanek (FH Campus Wien) erstellt.

<sup>2</sup> Die Frühjahrstagung des AK Soziale Dienstleistungen wurde von Olaf Lobermeier und Rainer Strobl (proVal), Stefanie Reiter (DJI) und Stefan Schmidt (schmidt evaluation) organisiert. Die letztgenannten drei Personen bilden zugleich das Sprecherteam des AK. Nicht nur für die inhaltliche Unterstützung und die Organisation der Tagungsräume, sondern auch für die Zubereitung eines köstlichen Chilis gebührt Olaf Lobermeier ein besonderer Dank. Herzlich danken möchten wir auch Jana Klemm (proVal) für ihre organisatorische Unterstützung während der Tagung, Edith Halves (HAW) für die Unterstützung bei der Arbeitsgruppenmoderation sowie unserem Kooperationspartner, dem Zentrum Demokratische Bildung in Wolfsburg, der uns die Tagungsräume zur Verfügung gestellt hat. Ein herzliches Dankeschön geht schließlich an Florentina Astleithner (FH Campus Wien), Benjamin Froncek (FernUniversität in Hagen) und Peter Stepanek (FH Campus Wien), die sich als Vortragende auf das Tagungskonzept eingelassen und während der gesamten Tagung engagiert mitgearbeitet haben.

## Berichte aus der Evaluationspraxis zu Herausforderungen von Wirkungsanalysen

Den Einstieg der Tagung bildeten zwei **Impulsreferate**, die Herausforderungen und Probleme bei der Wirkungsanalyse anhand von konkreten Beispielen aus der Evaluationspraxis darstellten.

Im ersten Vortrag „*Hat das Training Ihren Erwartungen entsprochen?*“. *Herausforderungen und Probleme bei der Wirkungsanalyse im Kontext von Diversity Trainings*“ (Autor/innen: Benjamin Froncek, Verena Piper, Dr. Agostino Mazziotta und Prof. Dr. Anette Rohmann, FernUniversität Hagen) wurden Herausforderungen bei der Erfassung von Wirkungen im Kontext von Maßnahmen zur Förderung interkultureller Kompetenzen betrachtet. Benjamin Froncek stellte vor, dass evidenzbasierte Praxis auch im Rahmen von Diversity Trainings zusehends gefordert wird.

Neben wenigen veröffentlichten Evaluationsstudien, die sich einzelnen Fragestellungen von Diversity-Trainings widmen, blieb bislang jedoch die Frage offen, wie sich die derzeitige Evaluationspraxis im Feld von Diversity-Trainings konkret ausgestaltet. Um diese Frage beantworten zu können, wurden Trainerinnen und Trainer von Maßnahmen zur Förderung interkultureller Kompetenzen (N = 172) im deutschsprachigen Raum in einem Online-Survey zu ihrer Evaluationspraxis befragt. Dabei zeigte sich, dass Trainerinnen und Trainer der Wirksamkeitsanalyse bei der Evaluation von Trainings durchaus eine hohe Bedeutung beimessen. Das Ausmaß des Einsatzes von Wirksamkeitsanalysen und die Art der methodischen Umsetzung spiegeln in der Praxis diese Bedeutung nicht immer entsprechend wider: Als Maß für den Trainingserfolg ist die Zufriedenheit der Trainingsteilnehmenden im Fokus der Evaluationen. Die Trainings werden überwiegend selbst evaluiert. Es kommen meist schriftliche Befragungen mit Selbstaussagen und einem Messzeitpunkt – im Anschluss an ein Training – zum Einsatz. Ein solches Vorgehen ist ggf. nützlich für die Qualitätssicherung, aber wenig aussagekräftig für die Wirksamkeitsüberprüfung. Der Survey zeigte also eine sichtbare Diskrepanz auf zwischen dem Anspruch, Wirkungen nachweisen zu wollen auf der einen und der berichteten Evaluationspraxis auf der anderen Seite.

Als mögliche Begründung dieser Diskrepanz wurden Herausforderungen bei der Umsetzung von Wirksamkeitsanalysen auf verschiedenen Dimensionen diskutiert: Erschwerende Bedingungen für den Einsatz von Wirkungsanalysen können unter anderem durch *Kontextfaktoren*, wie die Art und die organisationale Einbindung des Trainings, aber auch durch zur Verfügung stehende Ressourcen entstehen. Vonseiten der *Auftraggebenden* beeinflusst möglicherweise das vorhandene Wissen über Wirksamkeitsanalysen, ob Wirksamkeitsnachweise tatsächlich eingefordert werden. Aufseiten der *Trainerinnen und Trainer* stellt sich diesbezüglich die Frage nach der Haltung gegenüber evidenzbasierter Praxis sowie nach den Kompetenzen, Wirksamkeitsanalysen beurteilen oder gegebenenfalls selbst durchführen zu können. In Hinblick auf die *Teilnehmenden* eines Trainings spielen Aspekte der zeitlichen, kognitiven und motivationalen Beanspruchung für die Durchführung von Wirksamkeitsanalysen eine Rolle. Insbesondere im Zusammenspiel der aufgeführten Dimensionen ergibt sich für Evaluatorinnen und Evaluatoren die Herausforderung, ein kontextsensibles Design zu finden, das trotz der gegebenen Umstände in der Lage ist, Wirkungsdimensionen angemessen zu operationalisieren und objektive multiperspektivische Instrumente und Maße zu verwenden, um (nachhaltige) Wirkungen auf das Training zurückführen zu können.

Im zweiten Impulsvortrag „Zur Produktivität knapper Ressourcen in der Verbindung von Lehre – Praxis – Forschung: Wirkungsanalysen auf den Weg bringen“ (Autor/in: Florentina Astleithner und Peter Stepanek, FH Campus Wien) erläuterte Florentina Astleithner anhand von Beispielen aus der eigenen Evaluationspraxis, dass sich methodische und praktische Herausforderungen auch bereits vor der eigentlichen Durchführung von Wirkungsstudien stellen: Neben der Gestaltung des konkreten Forschungsdesigns rückt – auch aufgrund von steigenden Forderungen nach Wirkungsanalysen und Modellen der wirkungsorientierten Steuerung seitens verschiedener Stakeholder – die Frage der Finanzierung einer umfassenden Wirkungsstudie für Evaluierende sowie für Auftraggebende zusehends in den Fokus. Knappe Ressourcen tragen dazu bei, dass sich Organisationen als Träger von Programmen und Maßnahmen im Zuge der Vorbereitung von Wirkungsanalysen vermehrt mit diesem Anliegen auch an Universitäten und Fachhochschulen wenden. Die dadurch entstehende besondere Situation der Universitäten und Fachhochschulen, an der Schnittstelle zwischen Lehre und Forschung unter Einbindung von Studierenden und in enger Kooperation mit der Praxis Forschungsprojekte initiieren zu können, birgt innovative Ideen und hohe Nutzenpotenziale für alle Beteiligten (insbesondere Lehrende, Studierende und Praxiseinrichtungen) und wirft zugleich zusätzliche Herausforderungen auf, welche im vorgestellten Beitrag kritisch hinterfragt wurden: Unter anderem anhand eines Kooperationsprojekts zwischen einem Studiengang (Sozialwirtschaft und Soziale Arbeit) und einem Gemeinwesenprojekt in Wien wurde aufgezeigt, wie Masterarbeiten, begleitende Lehrveranstaltungen und Workshops zur Exploration des Forschungsfeldes als Vorbereitung einer Wirkungsstudie beitragen können. Basierend auf den vorliegenden Erfahrungen wurde diskutiert, inwieweit eine solche Kooperation tatsächlich die vielfältigen Ziele erreichen kann und z. B. sinnvoll sowie lehrreich für Studierende ist, das gewünschte Ergebnis für die Praxiseinrichtung liefert und aus Sicht der Forschungsverantwortlichen mit Blick auf fachliche Standards vertretbar ist. Außerdem wurde in diesem Kontext die Frage aufgeworfen, ob Kooperationen dieser Art eine Basis für die Durchführung eines umfassenden Auftragsforschungsprojektes bieten können bzw. wo die Grenzen (ökonomisch, methodisch, organisatorisch) liegen. In der bisherigen Umsetzung des Kooperationsprojekts erwies es sich als essenziell, Komplexität im Laufe des Forschungsprozesses aufzubauen, indem unter anderem unterschiedliche Perspektiven berücksichtigt, das Design theoriegeleitet und methodisch durchdacht und vorab auch geklärt wurde, welche Folgen in der Praxiseinrichtung entstehen können, wenn Wirkungen systematisch in den Blick genommen werden. Zugleich war die schrittweise Reduktion von Komplexität (beispielsweise im Sinne einer Klärung methodischer und ressourcenbedingter Grenzen sowie der Abstimmung von Rollen und Forschungs-, Beratungs- und Lehrendenperspektiven) ein wesentlicher Erfolgsfaktor für einen für alle Beteiligten gewinnbringenden Prozess. Der Ansatz erwies sich nach den Erfahrungen von Florentina Astleithner und Peter Stepanek insgesamt als gut geeignet, um verschiedene Perspektiven ineinander zu verschränken sowie Zeit und Raum für Auseinandersetzungen zu schaffen, um den Boden für Wirkungsanalysen vorbereiten zu können.

In drei **Arbeitsgruppen** ging es im Anschluss an die Impulsvorträge darum, pragmatische und – aus rein wissenschaftlicher Sicht vielleicht suboptimale Lösungen – für die effektive Durchführung von Wirkungsanalysen unter den Bedingungen knapper Ressourcen, eines engen Zeitrahmens und einer möglichst geringen Störung der Praxis zu diskutieren. Ein wichtiger Aspekt war dabei die kaum disku-

tierte Frage, ob die gefundenen Lösungen – auch mit Blick auf die Standards der DeGEval – noch vertretbar sind oder ob sie die Qualität der Evaluationsergebnisse gefährden. In diesem Zusammenhang konzentrierte sich die AG 1 auf methodische Herausforderungen bei einem primär quantitativen Vorgehen<sup>3</sup>, die AG 2 richtete ihr Augenmerk auf methodische Herausforderungen bei einem vorwiegend qualitativen Vorgehen und die AG 3 befasste sich mit methodischen Herausforderungen bei der Kombination von quantitativen und qualitativen Methoden. Den Arbeitsgruppen war es freigestellt, ob sie an die Eingangsreferate anknüpfen oder andere Probleme besprechen wollten. Die Ergebnisse der Arbeitsgruppen wurden anschließend im Plenum vorgestellt, zusammengeführt und diskutiert.

### **Ergebnisse der Arbeitsgruppen**

In den Arbeitsgruppen wurde durchaus mit Bezug auf die Eingangsvorträge anhand von Evaluationsbeispielen aus dem Kreis der Teilnehmenden diskutiert, wie Wirkungen unter den oft schwierigen Rahmenbedingungen einer Evaluation seriös nachgewiesen werden können. Die eigentliche Schwierigkeit besteht nach den Ergebnissen der Diskussion darin, die Wirkungen auf die realisierten Maßnahmen zurückzuführen. Das Problem der knappen Ressourcen stellt sich dabei aus Sicht der Tagungsteilnehmenden oftmals sowohl im Vorfeld von Wirkungsuntersuchungen als auch in allen Phasen des eigentlichen Evaluationsprozesses. Die Teilnehmenden diskutierten verschiedene spezifische Probleme und Lösungswege in den einzelnen Phasen. Nachstehend sind die zentralen Ergebnisse wiedergegeben.

#### *Planungsphase:*

Bei der Konzeption und Planung einer Evaluation werden unter anderem die grundlegenden Weichen für eine effektive Erhebung gestellt. Insbesondere bei qualitativen Erhebungen muss den befragten Personen ein Rahmen ermöglicht werden, in dem sie ungestört, ohne Zeitdruck und frei von sozialer Erwünschtheit ihr Wissen und ihre Haltungen zu den jeweiligen Fragestellungen Ausdruck verleihen können. Wer bereits bei der Planung Ressourcen sparen möchte, läuft große Gefahr, keine bzw. kaum verwertbaren Ergebnisse zu generieren. Gerade am Anfang können die Voraussetzungen für ein ressourcenarmes Arbeiten geschaffen werden. Als wesentliche „Stellschrauben“ wurden die Zielexplikation mit den Auftraggebenden und die Wahl des Designs, aber auch die Planung der Erhebung genannt:

- *Klärung des Zwecks/Ziels der Erhebung, der Nutzenerwartungen und der Fragestellungen bzw. Hypothesen mit den Auftraggebenden:* Als Voraussetzung für eine gelingende Wirkungsanalyse wurden hinreichend konkrete Ziele betrachtet, die gemeinsam mit den Auftraggebenden festgelegt werden. Dabei sollte man diese – auch in Abwägung mit den zur Verfügung stehenden Ressourcen – auf das Wichtigste reduzieren, bspw. nur den Zweck „Optimierung“ verfolgen und nur

---

<sup>3</sup> Nach Erörterung der von den Anwesenden praktizierten Erhebungsmethoden bzw. -quellen beschränkte sich AG die Diskussion auf die aus der Erfahrung der Anwesenden in der Evaluation am häufigsten angewendeten qualitativen Methoden: Interview und Gruppendiskussion. Weitere Methoden wie Beobachtung und Skype-Interviews wurden in der Diskussion nur am Rande besprochen.

zwei Fragestellungen/Hypothesen untersuchen. In dieser frühen Phase ist es besonders wichtig, zu prüfen, welches Design für welche Zielstellung angemessen ist und realisiert werden kann und welche Aussagen zur Wirksamkeit sich daraus ableiten lassen bzw. welche Grenzen ggf. bestehen. Die Auftraggebenden sollten bereits in dieser Phase dafür sensibilisiert werden, welche Schwierigkeiten darin bestehen, Wirkungen auf die realisierten Maßnahmen zurückzuführen, welche Aussagen sich bei welchen Design(-alternativen) generieren lassen und welcher Wirkungsbegriff angemessen ist (Erwartungsmanagement). Kritisch wurde insbesondere die Tendenz einer weitverbreiteten, unzureichenden Unterscheidung zwischen Zufriedenheitsmessungen und Wirkungsnachweisen gesehen. Gerade im Bereich der Evaluation sozialer Dienstleistungen muss (selbst-)kritisch geprüft werden, ob der Nachweis „knallharter Fakten“ durch quasi-experimentelle Designs realisierbar und im Einzelfall gegenstandsangemessen ist. Dabei sollten sich Evaluatorinnen und Evaluatoren auch vor Augen halten, dass Auftraggebende oftmals wiederum Legitimationszwängen von außen unterliegen. Diese zweite (Erwartungs-)Ebene kann dadurch berücksichtigt werden, dass bspw. den Auftraggebenden Argumentationshilfen für den Umgang mit Anforderungen und Nachfragen von außen zur Verfügung gestellt und Machbarkeitsgrenzen offen kommuniziert werden.

- *Wahl des Designs:* In diesem Zusammenhang nahm, insbesondere in der Arbeitsgruppe zu quantitativen Verfahren, die Diskussion über die Bildung von Kontrollgruppen einen breiten Raum ein. Die Teilnehmenden waren sich einig, dass ein sogenanntes RCT-Design (eine randomisierte, kontrollierte Studie) im Bereich sozialer Dienstleistungen kaum umsetzbar ist. In diesem Bereich werde stattdessen häufig argumentiert, dass der Ausschluss der Kontrollgruppenmitglieder von der Teilnahme an einem Angebot ein ethisches Problem darstellt. Als realisierbare Möglichkeit wurden in diesem Zusammenhang sogenannte Warte-Kontrollgruppen genannt. In diesem Fall nehmen die Mitglieder der Warte-Kontrollgruppe zu einem späteren Zeitpunkt an den Maßnahmen teil. Anspruchsvoller wäre ein sogenanntes Crossover-Design, bei dem zwei Gruppen an zwei Maßnahmen in unterschiedlicher Reihenfolge teilnahmen. Wenn es gelingt, Zeitreihen zu erfassen, dann kann einem Diskussionsbeitrag zufolge unter Umständen auch ganz auf eine Kontrollgruppe verzichtet werden. Unter Umständen sei es auch möglich, eine ausgearbeitete Programmtheorie, bspw. mit einem Strukturgleichungsmodell, zu überprüfen. Allerdings würden hierfür dann wieder erhebliche Ressourcen benötigt.
- *Planung der Erhebung:* Eine sorgfältige Planung einer Erhebungssituation (bspw. eine gelungene Terminierung von telefonischen Interviews und Vereinbarung des benötigten Zeitfensters, kriteriengeleitete Zusammenstellung der Personen für Gruppendiskussionen, ungestörte Räumlichkeiten etc.) ist essenziell. Es wurde angeregt, dass Evaluatorinnen und Evaluatoren prüfen sollten, ob Instrumente mit „Doppelfunktion“ (Nutzen für das Projekt/Programm und die Evaluation) einsetzbar sind. Ist eine Erhebung verzichtbar, wenn z. B. Memos von Mitarbeitenden und Teilnehmenden eines Projekts zu bestimmten Projektaktivitäten erstellt werden? Für die Projektbeteiligten kann das der Reflexion dienen, und das Evaluationsteam kann die Memos in einer Dokumentenanalyse auswerten. Ein weiteres Beispiel ist das Beschreiben und Auswerten von Papiertischdecken oder Flipcharts von Kleingruppen bei Großveranstaltungen, z. B. im Rahmen von O-

pen-Space-Veranstaltungen. Zudem wurde überlegt, Informationen, die in Neuen Medien generiert werden und die ohnehin verfügbar sind, als Datengrundlage mit einzubeziehen.

- *Instrumentenentwicklung:* Aus Diskussionen der Teilnehmenden ging hervor, dass der Leitfaden/die Frageroute erst nach der Klärung des Zwecks/Ziels, der Fragestellungen, der Stichprobe, der Erhebungssituation und der Auswertungsmethode entwickelt werden sollte. Den Auftraggebenden sollte transparent gemacht werden, welche Fragen zur Untersuchung welcher Fragestellungen und Hypothesen gestellt werden. Das vermeidet bzw. reduziert Missverständnisse und Erklärungsbedarfe.

#### *Erhebungsphase:*

Für die Erhebungsphase wurde in den Arbeitsgruppen diskutiert, welche ressourcenschonenden Zugänge gefunden und welche Personengruppen in die Erhebung einbezogen werden können.

- *Zugang über Communities:* Um den Aufwand zur Ansprache von Probanden und die Erschließung von Zugängen zur Zielgruppe zu verringern, wurde in einer Arbeitsgruppe vorgeschlagen, Personen und Communities als „Türöffner“ zu nutzen, die ohnehin bereits Zugang zu den anvisierten Erhebungspersonen haben (z. B. Sozialarbeiter/innen).
- *Einbeziehung der Praxis in die Erhebung:* Diskutiert wurde, ob es sinnvoll sein könnte, Praktiker für die Durchführung von Wirkungsanalysen zu qualifizieren, beispielsweise indem entsprechende Inhalte als Module in die Ausbildung von Sozialarbeiterinnen/-arbeitern und Multiplikatorinnen/Multiplikatoren eingespeist werden. Hier wurde jedoch tendenziell – vor allem in Abhängigkeit von den eingesetzten Methoden und Themen/Zielen der Evaluation – eine (zeitliche und fachliche) Überforderung der Praxis gesehen. Praktiker/innen könnten dennoch systematisch auf Wirkungsindikatoren achten und Hinweise für Wirkungen protokollieren. Selbstbeobachtung und Reflexion könnten dabei in Maßnahmen und Programmen als Bestandteil von Kurskonzepten/Schulungsinhalten etc. noch stärker integriert werden. In der Arbeitsgruppe zu Mixed Methods wurde angeregt, Erhebungsdaten, die von Praktikerinnen und Praktikern generiert werden, mit Expertenerhebungen zu triangulieren.
- *Erhebung durch Peers:* Ferner wurde vorgeschlagen, beispielsweise bei der Evaluation von Bildungsmaßnahmen für Jugendliche, Peers in die Erhebung einzubeziehen. Aufgrund fehlender Erfahrungswerte wurde dieser Punkt nicht abschließend diskutiert.

#### *Dokumentations- und Auswertungsphase:*

Für die Phase der Ergebnisdokumentation, Analyse und Aufbereitung wurden folgende Ideen für eine ressourcenschonende Umsetzung gesammelt und mit Blick auf die Einhaltung von fachlichen Standards erörtert.

- *Dokumentation:* Der Spielraum für die Dokumentation ist zum einen abhängig vom Zweck/Ziel der Erhebung und zum anderen davon, wie die Erhebungsergebnisse genutzt werden sollen. Auch diesbezüglich ist eine Klärung mit den Auftraggebenden anzuraten (siehe Planungsphase). Gerade

beim Einsatz qualitativer Verfahren stellt sich die Frage nach dem Umfang der Dokumentation. Die Spanne kann durch zwei Beispiele verdeutlicht werden:

- Sollen die Ergebnisse „nur“ dazu genutzt werden, um mit einem Projektteam Optimierungsmöglichkeiten für die Durchführung einer zweistündigen Aktivität zu erarbeiten, reicht ggf. ein Protokoll mit den zentralen Aussagen der Befragten, die in einem Workshop mit dem Projektteam diskutiert und interpretiert werden.
- Sollen die Ergebnisse genutzt werden, um die Wirkungen eines umfangreichen und kostenintensiven Programms zu beurteilen, damit Auftraggebende über dessen Fortführung entscheiden können, sind vollständige Transkripte (von mehreren Erhebungen zur Realisierung einer Datentriangulation) zu empfehlen, die von mindestens zwei Forscherinnen bzw. Forschern getrennt voneinander ausgewertet werden. Die Analyseergebnisse werden dann anschließend zusammengeführt (Forschertriangulation).

In der AG zu qualitativen Verfahren wurde zudem diskutiert, dass bestimmte Auswertungsmethoden nach spezifischen Dokumentationen verlangen. Beispielsweise erfordert die dokumentarische Methode nach Bohnsack ein Volltranskript mit Kennzeichnungen zur Lautstärke, Betonung, Unterbrechungen etc. Die Evaluationspraxis der Mitdiskutierenden wird den Erfahrungsberichten zufolge aber wesentlich geprägt von der inhaltsanalytischen Methode nach Mayring. Nach den Erfahrungen der Mitdiskutierenden sind die folgenden zwei Fälle am häufigsten anzutreffen und folgende Vorgehensweisen vertretbar:

- Interview: Das Telefonat wird auf Tonband mitgeschnitten. Der/die Interviewer/in fertigt während des Interviews stichwortartige Notizen des Verlaufs sowie der zentralen Aussagen an und erstellt anschließend ein Gedächtnisprotokoll. Abschließend wird das Tonband komplett abgehört, die für die untersuchten Fragestellungen wichtigen Aussagen werden ausführlich in eigenen Worten wiedergegeben oder vollständig transkribiert und damit das Gedächtnisprotokoll überarbeitet, ergänzt sowie korrigiert. Wörtliche Zitate werden immer vollständig transkribiert.
- Gruppendiskussion: Hier empfiehlt sich nach der Erfahrung der Anwesenden eine zweite Person zur Protokollführung. Die Moderation einer Gruppenerhebung erfordert eine hohe Aufmerksamkeit und Beobachtung auch nonverbaler Kommunikation in der Gruppe, was kaum zulässt, nebenbei Ergebnisnotizen anzufertigen. Die Dokumentation im Anschluss an die Erhebung entspricht denen des Interviews (siehe oben).
- *Auswertung:* Bezüglich der Auswertungsmethoden wurden in der AG qualitative Verfahren Erfahrungen bei der Anwendung der inhaltsanalytischen Auswertung nach Mayring und der dokumentarischen Methode nach Bohnsack ausgetauscht. Die am weitesten verbreitete Software zur Textanalyse ist MAXQDA, zur Transkription f4.

### *Phase der Ergebnisaufbereitung und Rückspiegelung der Evaluationsergebnisse:*

In den Arbeitsgruppen wurden ferner einige Anregungen für eine ressourcenschonende Aufbereitung und Weitergabe von Resultaten gesammelt.

- *Ergebnisaufbereitung:* Bezüglich der Aufbereitung von Evaluationsergebnissen für die Auftraggebenden und ggf. weitere Personengruppen wurden als Alternativen für einen ausführlichen schriftlichen Bericht Präsentationen sowie Schaubilder an Moderationswänden mit unterschiedlichen Visualisierungsmethoden angesprochen.
- *Rückspiegelung der Ergebnisse:* Von Teilnehmenden wurde es als relevant erachtet, den Mut zu haben, „kleine“ Ergebnisse darzustellen und zu erklären, wie die Qualität und Aussagekraft von Ergebnissen ist. Wie bereits in der Planungsphase ist es vonseiten der Evaluatorinnen und Evaluatoren sinnvoll, wiederum auch die Anforderungen von außen an die Auftraggebenden (bspw. aus dem politischen Raum) zu reflektieren und Argumente für die Weitergabe der Ergebnisse an zusätzliche Stakeholder zu liefern. Auch hier gilt es, Sensibilisierungsarbeit zu leisten und gegenüber Auftraggebern, weiteren Stakeholdern – und auch anderen Evaluatorinnen und Evaluatoren – zu klären, dass harte Wirkungsnachweise nicht immer erbracht werden können und ggf. auch „kleine“ Ergebnisse für die Legitimation von Maßnahmen und Programmen ausreichen. Als Idee wurde zudem geäußert, neben dem Evaluationsteam weitere Personengruppen in die Kommunikation von Ergebnissen einzubeziehen, um eine Ergebnisbetrachtung aus verschiedenen Perspektiven zu ermöglichen und das Verständnis für die Ergebnisinhalte abzusichern.

Für eine ausführliche Erörterung der Themen Auswertung und Aufbereitung der Evaluationsergebnisse fehlte in den Arbeitsgruppen leider die Zeit.

### **Fazit**

Die Idee, zu dem vielleicht etwas unüblichen Tagungskonzept mit einem hohen Anteil an Arbeitsgruppen, entstand während der DeGEval-Jahrestagung 2014 in Zürich. Dort hatte das AK-Sprecherteam den Eindruck, dass es oft weniger die großen Themen als vielmehr die alltäglichen Probleme im Evaluationsalltag sind, die Evaluatorinnen und Evaluatoren unter den Nägeln brennen. Die Tagung war in gewisser Weise ein Experiment, das nur gelingen konnte, wenn die Teilnehmenden bereit waren, in den Arbeitsgruppen auch ein wenig „aus dem Nähkästchen zu plaudern“, Dinge zu fragen und Lösungen vorzuschlagen, die sie bisher eher für sich behalten hatten.

Rückblickend war das Konzept ein voller Erfolg. Das große Interesse im Vorfeld der Tagung zeigte, dass das AK-Sprecherteam mit dem Tagungsthema und dem Tagungsformat einen Nerv getroffen hatte. Nach den Eingangsreferaten, die den Rahmen der Tagung noch einmal verdeutlichten, ist es in den Arbeitsgruppen gelungen, eine vertrauensvolle und offene Diskussionsatmosphäre zu schaffen, in der neue und ungewöhnliche Ideen vorgestellt und diskutiert wurden. Die überaus positiven Rückmeldungen am Ende der Tagung bestätigen die Einschätzung, dass es einen erheblichen Bedarf an einem Tagungsformat gibt, das einen offenen und vertrauensvollen Austausch fördert. Zudem ist dieses Format mit einem relativ begrenzten Aufwand zu realisieren. Das AK-Sprecherteam kann es daher als Ergänzung zur üblichen Form einer Fachtagung uneingeschränkt empfehlen.