

Wolfgang Beywl/Lars Balzer

Aufbau von Evaluationskompetenzen für interne Schulevaluation durch projektbezogene Fortbildung

Zusammenfassung

Interne Evaluationen werden aktuell wieder stärker diskutiert. Der Beitrag definiert zunächst schulinterne Evaluation und identifiziert Varianten nach Größe des Evaluationsgegenstandes und nach der Bearbeitungsweise. Für drei Typen von Varianten werden erforderliche Kompetenzen skizziert. Das Kompetenzportfolio für schulinterne Evaluationen lässt sich inhaltlich in zehn Schritte gliedern, nach denen Evaluationen geplant und durchgeführt werden können. Es wird abschließend beschrieben, wie diese Kompetenzen in einer Evaluationsfortbildung entlang der zehn Schritte projektbezogen erarbeitet und angewendet werden können.

Schlüsselwörter: Selbstevaluation, interne Evaluation, Schulevaluation, Unterrichtsevaluation, Unterrichtsentwicklung, Schulentwicklung

Evaluation Capacity Building for Internal School Evaluation by Project-centered Training

Summary

Internal evaluations are currently being discussed more intensively. The paper initially defines internal school evaluation and identifies variants, depending on the size of the evaluand and the treatment mode. Required competencies are outlined for three types of variants. The competence portfolio for internal school evaluations can be broken down to ten steps, which guide evaluation planning. The paper outlines how the skills can be acquired in an evaluation training alongside the ten steps, which includes a real evaluation project.

Keywords: self-evaluation, internal evaluation, school evaluation, evaluation of teaching, improvement of teaching, school development

1. Wichtigkeit, Seltenheit und neue Chancen

Interne Evaluation gilt vielfach als zentral für die schulische Qualitätsentwicklung, als wirkmächtig zur Entwicklung von Unterricht und Schule. Nevo (2009) erachtet die interne Evaluation als nachhaltige Investition, um das Wissensmanagement der Bildungsorganisation zu stärken, und auch als Chance, Steuerungsentscheidungen vom Zentrum an die Peripherie, von der zentralen Bildungsverwaltung an die lokale Schule zu verlagern. Die deutsche Kultusministerkonferenz fordert im Zusammenhang mit der Einführung von Bildungsstandards, „verstärkt Unterrichtsentwicklung zu betreiben, [...] auch, sich regelmäßig des Erfolgs der Arbeit zu vergewissern (interne Evaluation)“ (KMK 2005, S. 10). Pietsch (2011) weist aus, dass interne Evaluation in den Schulgesetzen der meisten Bundesländer eine prominente Rolle einnimmt. Für die Deutschschweizer Kantone gilt dies in ähnlicher Weise (vgl. IDES 2007). Lehrbuchautoren und -autorinnen betonen, wie wichtig interne Evaluation für die Qualität der Schule sei (vgl. den kritischen Überblick von Zenkel 2015).

Empirische Studien zeichnen ein enttäuschendes Bild in Bezug auf Akzeptanz und Nachhaltigkeit: Bei Schulleitenden der D-A-CH-Länder gehören Evaluationen und andere datenbezogene Aufgaben zu den bei weitem unbeliebtesten und belastendsten (vgl. Huber/Wolfgramm/Kilic 2013). Pietsch (2011, S. 13) konstatiert auf Basis mehrerer Quellen (PISA 2003, Syntheseberichte von Schulinspektionen/externer Schulevaluation, GEW-TALIS), dass schulinterne Evaluation „kaum entwickelt wurde und derzeit nicht zu den Stärken schulischer Arbeit gehört“. Dieser Befund gilt auch für die USA, das „Mutterland“ der Schulevaluation. Hier bleibe eine entwicklungsorientierte schulische Programmevaluation stets sekundär gegenüber den stark verbreiteten standardisierten landesweiten Leistungsmessungen (vgl. King/Rohmer-Hirt 2011, S. 76). Gärtner (2015, S. 34) spricht auf Basis eigener empirischer Untersuchungen zu Brandenburg sowie einer Auswertung internationaler Studien von einem „geringen Anteil nutzbringender interner Evaluation“.

Während also *auf der konzeptionellen Ebene* gesetzliche Grundlagen, bildungstheoretische Annahmen und Befunde der Lehr-/Lernforschung dezidiert für interne Evaluation sprechen, so wird sie *in der Praxis* – wenn überhaupt – allzu oft als lästige Pflichtaufgabe abgearbeitet, da ihr kaum konkreter Nutzen für Schule und Unterricht zugetraut wird.

Eine zweite Chance für die interne Evaluation ergibt sich aus der – gerade im deutschsprachigen Raum – breit rezipierten Meta-Meta-Analyse von John Hattie (2014) zu den Einflussfaktoren auf schulische Lernleistungen. Eine seiner zentralen Schlussfolgerungen lautet, dass Lehrpersonen und Schulleitende zu Evaluierenden ihrer eigenen professionellen Tätigkeit werden sollen. Dies betrifft besonders den

starken Faktor „Formative Evaluation des Unterrichts“ (Hattie 2015, S. 215). Diese Forderung untermauert das von Unterrichts-, Schul- und Bildungsforschern (z.B. Andreas Helmke, Eckhard Klieme, Hilbert Meyer, Michael Schratz) vertretene Plädoyer für interne Evaluation. Es gilt, interne Evaluation konsequent auf die Verbesserung des Unterrichts und die Stärkung des Lernens (der Schülerinnen und Schüler wie der Lehrkräfte) auszurichten, also auf die Mikro-Ebene des Unterrichts und die Meso-Ebene der unterrichtsübergreifenden Kooperation. Derart lernorientiert Evaluierende sollten sich vor allem beim Aufbau von Evaluationsvermögen und Evaluationskultur in der Schule engagieren (vgl. Dahler-Larson 2009). Es geht um schulinternes Evaluationsvermögen – verstanden als Fähigkeit einer Bildungsorganisation und ihrer Mitglieder, Evaluationen zu planen, durchzuführen und zu nutzen, um daraus individuell und organisational zu lernen. Hierfür gibt es nach Cousins et al. (2014) zwei unabdingbare Voraussetzungen:

- ermöglichende und stützende „Strukturen“, d.h. Evaluationsaufgaben in Stellenbeschreibungen der Lehrenden, entsprechende Arbeitszeitkontingente, fixierte Arbeitsgruppen-/Unterrichtsteam-Termine zur Erarbeitung sowie schulweite Kommunikations- und Präsentationskanäle zur Vermittlung von Ergebnissen und zur Einleitung ihrer schulweiten Nutzung;
- Kompetenzen zur Planung und Umsetzung von Evaluationen, die sich Schulinterne aneignen.

Welches Verständnis von interner Evaluation dafür zugrunde zu legen ist und welche Kompetenzprofile je nach Gegenstand der Evaluation und dessen Größe erforderlich sind, behandelt der nachfolgende Abschnitt.

2. Verortung und Varianten interner Schulevaluation

Wie im einleitenden Beitrag von Böttcher und Hense (2016) in diesem Heft wird Evaluation hier als systematisch geplantes und umgesetztes, auf den jeweiligen Gegenstand maßgeschneidertes Beschreibungs- und Bewertungs-Verfahren verstanden, das

- auf spezifische Handlungsfolgen ausgerichtet ist,
- dabei Methoden der empirischen Sozialforschung einsetzt und
- Bewertungen nachvollziehbar auf Kriterien abstützt.

In Bezug auf den von ihr zu verfolgenden Zweck kann eine Evaluation unterschiedliche¹ Prioritäten setzen:

- Formative Evaluation soll zur optimalen „Ausformung“ des Evaluationsgegenstandes, z.B. eines bestimmten Lehr-Lernarrangements, beitragen. Sie soll die-

1 Diese Unterscheidung geht auf Michael Scriven (1972/1966) zurück.

sen stabilisieren oder verändern helfen und verfolgt primär den Zweck der Verbesserung.

- Summative Evaluation will einen „Summenstrich“, eine Bilanz zum Evaluationsgegenstand ziehen. Typische Zwecke sind die Rechenschaftslegung (z.B. gegenüber den Bildungsbehörden) und die Vorbereitung von Richtungsentscheidungen.

Im Bildungsbereich richten sich Evaluationen vielfach auf *Programme* („Unterrichtssequenzen“, „Kurse“, „Angebote“, „Projekte“ usw.; vgl. Abschnitt 4 in Böttcher/Hense 2016). Entsprechend tragen viele englischsprachige Standardwerke *Program Evaluation* im Titel. Programme basieren auf einem (möglichst schriftlichen) Konzept. In der Schule sind solche „*sichtbaren*“ *Programme* (wie z.B. eine Projektwoche oder die Berufswahlorientierung) die Ausnahme gegenüber dem Regelfall des *alltäglichen Unterrichts*. Dabei weisen beide ähnliche Grundelemente auf: Lernausgangslagen, die auf Ziele hin verändert werden sollen, sowie darauf zugeschnittene Methoden/Interventionen, die das Lernvermögen der Schülerinnen und Schüler mobilisieren und schließlich zu Lernresultaten führen sollen. Konzepte für Unterricht und damit das Kernelement der Programmförmigkeit sind zwar gegeben, doch – wie auch der Unterricht selbst – *kaum* für Dritte *sichtbar*. Der „normale“ Unterricht einerseits, spezifische schulische Angebote andererseits erfordern daher unterschiedliche Ansätze der Evaluation (siehe Abbildung unten). Dieser Beitrag konzentriert sich auf programm-/projektförmig abgegrenzte bzw. abgrenzbare mittelgroße Evaluationsgegenstände.

Programme im Bildungsbereich werden von *Organisationen* durchgeführt und (mit-)verantwortet. Aus dem für Organisationsförmigkeit konstitutiven Element der Mitgliedschaft (vgl. Kühl 2011) ergibt sich als weitere Unterscheidung die nach Evaluationsarten: zum einen der Fall, dass die Evaluation wesentlich durch Mitglieder² der Schule geplant und umgesetzt wird (interne Evaluation), und zum anderen der, dass dies Nicht-Mitglieder der Organisation tun (externe Evaluation).

Für eine bestimmte Evaluation sollte ein Schwerpunkt entweder auf formativ oder summativ gelegt werden. Hieraus ergeben sich Präferenzen für die eine oder andere Art der Durchführung: Die summative Rolle kann durch externe Evaluationsfachleute, die keine Interessenkonflikte haben, glaubwürdiger wahrgenommen werden. Hingegen wird Organisationsinternen ob ihrer genauen Kenntnisse der Schülerschaft, der sozialen Beziehungen und des sozialen Kontextes der Schule eher zugetraut, dass sie mit unmittelbarem Nutzen für die pädagogische Praxis formativ evaluieren können. Oft ist es sinnvoll, intern Evaluierende extern zu beraten, damit

2 Dies sind namentlich Lehrkräfte, Schulleitende, evtl. weitere Angestellte. Sie treten der Schule aus freiem Willen per Arbeitsvertrag bei und können ebenso austreten. Freiwilligkeit ist für Schülerinnen und Schüler zumindest während der Pflichtschuljahre nicht gegeben. Organisationstheoretisch gesehen sind sie keine Mitglieder; evaluationstheoretisch zählen sie zu den Stakeholdern (s.u. Abschnitt 4, 2. Schritt).

sie blinde Flecken aufdecken oder sich aus ihrer Organisationsmitgliedschaft ergebende Herausforderungen besser bewältigen (tatsächliche oder vermeintliche pädagogische Wertkonflikte, Widerstände von Kolleginnen und Kollegen, die Autonomie- bzw. Kontrollverlust befürchten, u.v.m.).

Auf Makroebene, z.B. bei der Kooperation von Grund- mit weiterführenden Schulen, wird besser extern evaluiert, denn die jeweils Internen geraten allzu schnell in Interessenkonflikte. Die Mikro-Ebene des Unterrichts gehört hingegen in die Domäne der internen Evaluation. Dazwischen, auf der Meso-Ebene, sind beide Hauptarten geeignet, die interne allerdings eher, wenn es formativ um Verbesserung geht.

Interne Evaluation wird im Bildungsbereich oft als „Selbstevaluation“ bezeichnet (vgl. Buhren 2007; Burkard/Eikenbusch 2000; von Saldern 2010), deren Pendant als „Fremdevaluation“.³ Im Unterschied zur durch (Nicht-)Mitgliedschaft leicht operationalisierbaren Gegenüberstellung extern/intern ist die Unterscheidung fremd/selbst komplex und basiert auf dem Konzept der *Kultur*. Diese Differenz kann sich aus (Professions-)Kulturen und damit verbundenen Werten von Evaluierenden einerseits, von für das pädagogische Programm verantwortlichen Stakeholdern andererseits herleiten.⁴

Wie Abbildung 1 zu entnehmen, gibt es viele Varianten schulinterner Evaluation, je nach der *Größe der zu bearbeitenden Evaluationsgegenstände* und je nach gewählter *Bearbeitungsweise*:

Bezüglich *Größe* des Gegenstands gibt es folgende Pole:

- fokussierte Evaluationen, bezogen auf einen pädagogischen Gegenstand z.B. im Umfang weniger Unterrichtsstunden;
- ausgedehnte Evaluationen zu schulweit oder gar übergreifend relevanten Fragestellungen, die z.B. aus dem Schulprogramm abgeleitet oder durch die Ergebnisse externer Evaluationen/Inspektionen ausgelöst sind.

Besonders letztgenannte richten sich oft auf die Schule als Organisation. Zwischen diesen Polen liegen zahlreiche Evaluationsgegenstände unterschiedlichen Umfangs.

3 Wissenschaftstheoretisch präziser (vgl. Eckensberger 2015) werden „etische“ (fremd/generalisierend/standardisiert) von „emischen“ (selbst/ideographisch/responsiv) Zugängen unterschieden. Ethisch wird der Gegenstand von außerhalb des Systems evaluiert, mit feststehenden Fragestellungen und allgemein gültigen Bewertungskriterien, auch um ihn mit anderen zu vergleichen. Emisch wird von innen her evaluiert, basierend auf systemspezifischen Fragestellungen und Kriterien. Angesichts solch spannungsgeladener Themen wie Koedukation, Inklusion oder leistungsdifferenzierter Klassenbildung ist offensichtlich, welchen Unterschied ein etischer oder emischer Evaluationszugang macht. Dies ist besonders in der deutschsprachigen Forschung zur Schulevaluation kaum theoriebasiert thematisiert.

4 Als „fremd“ wird z.B. von pädagogischen Fachpersonen wahrgenommen, wenn ihre Angebote durch eine von Juristen oder Betriebswirtinnen geführte Stelle wie etwa ein Rechnungsprüfungsamt evaluiert werden.

Einige von diesen sind in Spalte 3 der Abb. 1 aufgeführt. Die Größe eines Evaluationsgegenstandes ergibt sich aus seiner zeitlichen Dauer, der Anzahl der als Lehrende oder Lernende oder anderweitig Mitwirkenden, der Menge erforderlicher Abstimmungs- und Übergabepunkte u.a.

Die geeignete *Bearbeitungsweise* hängt mit der Größe des Evaluationsgegenstandes zusammen:

- Ein wenige Unterrichtsstunden umfassender, sehr fokussierter Evaluationsgegenstand kann von einer Lehrkraft (oder einem Tandem) selbständig und beiläufig im Unterricht evaluiert werden. Unterrichten und Untersuchen können *integriert*, in einem Zug durchgeführt werden. In Fortführung der von Schön (1983) für professionelles Lehren geprägten *reflection in action* könnte man von *evaluation in action* sprechen. Es gibt Nähe zur Aktionsforschung (vgl. Altrichter/Posch 2007) oder zur Selbstevaluation als Bestandteil methodischen Handelns (vgl. Bestvater/Beywl 2015). Diese Bearbeitungsweise wird „individuell/Tandem“ genannt.
- Größere interne Evaluationen müssen hingegen von speziell dafür zuständigen Lehrpersonen mit reduzierter Unterrichtsverpflichtung oder nicht unterrichtenden Evaluationsfachleuten *arbeitsteilig* bearbeitet werden. Diese tragen keine (Haupt-)Verantwortung für das zu evaluierende Bildungsangebot. In sehr großen Schulen gibt es evtl. eine eigene Stabsstelle „Evaluation“. Sie ist für die Planung und Durchführung der Evaluation zuständig, während (andere) Lehrpersonen den Unterricht/die pädagogischen Maßnahmen steuern. Als Interne sind sie dem Wert- und Normensystem der Schule verpflichtet, und das Kollegium verlangt von ihnen Loyalität. Sie evaluieren im Auftrag der Schulleitung und in Abstimmung mit weiteren Beteiligten. Diese Variante der internen Evaluation wird hier zur besseren Unterscheidbarkeit „Inhouse-Evaluation“ genannt. Oft werden diese Ansätze (missverständlich) als „Schulevaluation“ oder „institutionelle Selbstevaluation“ bezeichnet. Diese die pädagogische Arbeit „unterstützende“ Bearbeitungsweise wird im Anschluss an Zenkel (2015) als „auxiliar“ bezeichnet.
- Für Evaluationsgegenstände in dazwischen liegenden Größenordnungen eignet sich die kollegiale Bearbeitungsweise, angesiedelt zwischen individuell/Tandem und auxiliar. Hier arbeiten drei oder mehr Lehrkräfte zusammen, im Rahmen einer zumindest minimalen Projektorganisation.

Soweit an Schulen Unterrichtsteams, Fachschaften oder Qualitätsgruppen etabliert sind und es für diese verbindliche und anerkannte Ressourcen, Zeiten und Abläufe gibt, ist es möglich, kollegial zu evaluieren. Wenn jedoch diese strukturellen Voraussetzungen (vgl. Cousins et al. 2014) schwach sind, sind individuelle/Tandem-Ansätze und auxiliar angelegte Evaluationen leichter realisierbar.

Wie Abbildung 1 zu entnehmen ist, ergeben sich durch die Kombination der Merkmale *Bearbeitungsweise* (Unterarten der internen Schulevaluation) und *Größe des Evaluationsgegenstandes* zahlreiche Varianten. Diese können zu den Typen I bis

III zusammengefasst werden, wobei es Überschneidungsbereiche gibt. Die folgende Abbildung kann auch als Prüfschema verwendet werden:

- Welche Bearbeitungsweise passt zum Evaluationsgegenstand?
- Welche Evaluationsgegenstände sind mit gegebenen Ressourcen und Kompetenzen bearbeitbar?

Abb. 1: Varianten interner Schulevaluation

Evaluationsgegenstand			Bearbeitungsweise		
Größe	Inhalt	Beispiele	individuell/ Tandem	kollegial	auxiliar
ausgedehnt ↑	Organisation	Ganze Schule, z.B. Qualitätsmanagement, Schulführung, Schulklima, inner-/außerschulische Kooperation			Typ III
		Pädagogisches Gesamtprogramm/Bildungsplanung der Schule			
↓ fokussiert	Pädagogik	Schulisches Projekt (z.B. Gesundheitsförderung, Inklusion, digitale Medien, Berufswahlorientierung u.a.)			
		Gesamtunterricht für Klasse/Jahrgangsstufe/Schulart (z.B. über ein bis drei Jahre)			Typ II
		Fächerübergreifender Unterricht (z.B. Projektwoche)			
		Fachunterricht in Klassen einer Stufe (z.B. über mehrere Monate oder ein Jahr)			
		Fachunterricht in einer Klasse (z.B. über ein halbes Jahr)		Typ I	
		Unterrichtseinheit (z.B. sechs bis zwölf Unterrichtsstunden)			

Quelle: eigene Darstellung

3. Kompetenzerwerb für schulinterne Evaluation

Die Annahme, dass für interne Schulevaluation keine spezifischen Kompetenzen erforderlich seien, ist weit verbreitet. So wird z.B. unterstellt, diese würden durch die seit über zehn Jahren aufgebauten Online-Selbstevaluations-Portale⁵ überflüssig. Die teils kostenlos verwendbaren Fragebogen und andere „qualitativ hochwertige Instrumente“ (vgl. Gärtner 2015) könnten durch Evaluationslaien sicher angewendet werden (vgl. von Saldern 2010). Allerdings ist – wie auch Gärtner festhält – der erhoffte durchschlagende Erfolg in Bezug auf den Aufbau einer lebendigen internen Evaluationskultur in den allermeisten Schulen ausgeblieben.

Hingegen erfordert eine nachhaltige, für Unterricht und Schule nutzenstiftende Evaluation evaluationsspezifische Kompetenzen (vgl. Abschnitt 3 in Böttcher und Hense 2016 in diesem Heft sowie Cousins et al. 2014). In einer Expertenorganisation wie der Schule werden Evaluationsvermögen und -kultur dann entstehen können, wenn möglichst viele der dort tätigen Lehrexperthen und -expertinnen Evaluationskompetenz als gelebten Bestandteil ihrer Professionalität betrachten. Evaluationen vom Typ I sind insofern grundlegend. Die gesamte Bandbreite schulinterner Evaluation zu aktivieren, ist Aufgabe der Schulleitung. Ihr Engagement ist ausschlaggebend für den Erfolg (vgl. Hattie/Masters/Birch 2016). Hattie (2014, S. 183) betont als dafür wichtige Haltung: „Lehrpersonen/Schulleitende sind überzeugt, dass ihre fundamentale Aufgabe darin besteht, ihr Lehren und das Lernen und die Lernleistung der Schülerinnen und Schüler wirkungsorientiert zu evaluieren“. Es dürfte sehr wenige Fachkräfte geben, die vertiefte Expertise für alle drei der nachfolgend skizzierten Typen interner Evaluation aufweisen. Daher ist in der Regel eine arbeitsteilige Zuständigkeit angeraten.

Typ I Individuell-kollegiale unterrichtsintegrierte Evaluation

Evaluationskompetenz für Typ I ist eine Erweiterung vorhandener didaktischer Experten-Kompetenz, die in mehrjähriger praktischer Unterrichtstätigkeit ausgebaut worden ist. Datenerhebung und -auswertung werden möglichst ohne Zusatzaufwand in den Unterricht integriert. Die Untersuchungen sind so angelegt, dass sie für das Erreichen von Lernzielen, das eigene pädagogische Lernen und den Dialog mit den Lernenden über den Unterricht möglichst unmittelbar Nutzen stiften können. Dies meint Hattie (2014, S. 164) mit „formativer Evaluation des Unterrichts“. Doch auch diese „Nebentätigkeit“ (vgl. Böttcher/Hense 2016 in diesem Heft) erfordert Kompetenzaufbau und Einüben in die Evaluationspraxis. Der von über 300

5 Um nur einige zu nennen: Der Vorreiter, SEIS, ursprünglich von der Bertelsmann Stiftung gefördert, hat seinen Betrieb nach zwölf Jahren Ende 2015 eingestellt. Das Portal IQUES-Online ist in Deutschland und der Schweiz seit ca. zehn Jahren verbreitet und zeichnet sich u.a. dadurch aus, dass es zahlreiche Materialien zur Unterrichtsentwicklung bereitstellt. Mehrere Bundesländer finanzieren bzw. betreiben seit einigen Jahren eigene Online-Instrumentensammlungen.

Lehrkräften erprobte Ansatz LUUISE (Lehrpersonen unterrichten und untersuchen integriert, sichtbar und effektiv) ermöglicht dies in kollegial gestützter, projektförmi-ger schulinterner Weiterbildung (siehe Härrä 2015; URL: <http://tinyurl.com/Luuisse>; Zugriffsdatum: 11.04.2016). Für die nachhaltige Optimierung des Lehrerhandelns sind derart auf Sichtbarmachen der eigenen Wirksamkeit angelegte Fortbildungen be-sonders effektiv (vgl. Lipowsky/Rzejak 2015).

Typ II Kollegial-auxiliare unterrichtsübergreifende Evaluation

Zusätzlich zu den Kompetenzen von Typ I sind für Typ II Evaluationskompetenzen erforderlich, welche die Befähigung zur Planung und Umsetzung fokussierter Evaluationen bis hin zu mittelgroßen Projekten umfassen. Sie können in einschlägi-ger Weiterbildungen erworben und in gecoachten Evaluationsvorhaben erprobt wer-den. Dabei bietet es sich an, die Weiterbildung inhaltlich-thematisch nach einer für die Evaluationspraxis orientierenden Schrittfolge zu gliedern (vgl. Textkasten), so dass der Transfer in die Praxis unmittelbar anschließen kann.

Typ III Auxiliare Projekt-/Programmevaluation – Ebene Schule

Diese verlangt eine umfassende Evaluationskompetenz. Es geht um die Befähigung, adaptiv für den jeweiligen Evaluationsgegenstand und seinen Kontext die leiten-den Evaluationsfragestellungen sowie geeignete Evaluationsmodelle (vgl. Beywl 2011) auszuwählen und diese konzeptionell wie zeitlich in die Schulentwicklung einzupassen. Neben grundständigem didaktischen sind auch ausgeprägtes empiri-sches Untersuchungswissen und -können sowie Beratungskompetenz erforder-lich. Dieses Kompetenzprofil (vgl. Russ-Eft et al. 2008) kann in einem postgradua-len wissenschaftlichen (Weiterbildungs-)Studium erworben und in mehrjähriger Evaluationspraxis vertieft werden.

Nachfolgend wird geschildert, wie Kompetenzen für interne Schulevaluationen des Typs II berufsbegleitend erworben werden können. Um dies zu erreichen, wurde die nachfolgend skizzierte zweiteilige Evaluationsweiterbildung konzipiert und erprobt.

Im ersten Teil der Fortbildung wird das evaluative Grundlagenwissen und -können erarbeitet. Dabei sollen die Teilnehmenden die systematische 10er-Schrittfolge für Planung und Durchführung von Evaluationen (vgl. Textkasten) erläutern, begründen und auf ein eigenes Evaluationsprojekt planend anwenden können.

Im *ersten Schritt* wird der *Evaluationsgegenstand* bestimmt. Dafür wird eine kurze, für Außenstehende nachvollziehbare Erstbeschreibung mit allen zentralen Elementen erstellt. Oft gilt es, die durch den Evaluationsgegenstand verfolgten Ziele zu präzisieren und ihnen Interventionen/Methoden/Lehr-Lernarrangements, Sozialformen, Medien/Materialien u.a. zuzuordnen (vgl. den Beitrag von Giel 2016 in diesem Heft, S. 149-162).

Im *zweiten Schritt* werden die am Evaluationsgegenstand interessierten *Akteure* bestimmt. Deren Werte und Interessen bezüglich des Evaluationsgegenstands und ihre Informationsbedarfe an die Evaluation sind zu identifizieren.

Im *dritten Schritt* geht es um die Festlegung der *Evaluationszwecke*: Unterscheidbar sind entscheidungsorientierte, rechenschaftslegungsorientierte, verbesserungsorientierte und wissensgenerierende Evaluation. Die auf die Zwecke abgestimmten *Evaluationsfragestellungen* fassen in Worte, welche Informationen die vorgesehenen Nutzenden benötigen.

Bewertungskriterien als im Vorhinein ausgewiesene Referenzmaßstäbe sind Voraussetzung für jede systematische und faire Bewertung. Sie werden im *vierten Schritt* festgelegt und bieten eine Bezugsbasis, auf die sich Werturteile zum Evaluationsgegenstand stützen.

Im *fünften Schritt* planen die Evaluierenden das methodische Vorgehen. Es gilt, ein passendes *Erhebungsdesign* sowie angemessene *Erhebungsmethoden* auszuwählen und einen Datenerhebungsplan zu entwerfen.

Die empirische Phase der Evaluation beginnt im *sechsten Schritt* mit der *Durchführung der Erhebungen*. Dabei sind die logistische und technische Seite ebenso zu beachten wie der Umgang mit Störungen und Konflikten.

Auf die Erhebungen folgen im *siebten Schritt* die *Auswertung* und *Interpretation* der Daten. Diese werden komprimiert, verbal zusammengefasst sowie in Bezug zu den Fragestellungen in deutende, erklärende Sinnzusammenhänge gestellt und zu Schlussfolgerungen verdichtet. Bewertungen und *Bewertungssynthese* erfolgen in Referenz auf die Bewertungskriterien.

Damit eine Evaluation Nutzen für Schule und Unterricht entfalten kann, muss im *achten Schritt* über ihre Ergebnisse *berichtet* werden: mündlich oder schriftlich, in Workshops, auf Papier oder mittels elektronischer Medien.

Evaluationen lohnen sich dann, wenn sie für und durch die schulische Praxis *genutzt* werden, was im *neunten Schritt* eingeleitet wird. Es geht nicht allein um die Verwendung der *Ergebnisse*, sondern auch darum zu klären, was der *Evaluationsprozess* bereits an Nützlichem für die Bildungspraxis erbracht hat.

Im abschließenden *zehnten Schritt* kann die Evaluation selbst bewertet werden. Mit der *Evaluation der Evaluation* ist diese abgeschlossen, und der Evaluationszyklus schließt sich, um bei Bedarf erneut zu beginnen.

Zwei didaktische Prinzipien leiten diesen ersten Teil der Weiterbildung: Mit dem „umgedrehten Unterricht“, dem *flipped classroom* (vgl. Bergmann/Sams 2012), wird die übliche Aufteilung dessen, was innerhalb (Wissensvermittlung) und außerhalb (Transfer/Anwendung) des Schulungsraumes stattfindet, „umgedreht“. Die Teilnehmenden erarbeiten sich das Evaluationswissen im Vorfeld und außerhalb der Präsenzphasen – selbstgesteuert, asynchron, im eigenen Lerntempo und ortsunabhängig. Dies geschieht durch die Auseinandersetzung mit dem an der 10er-Schrittfolge orientierten Lehrtext mit kapitelweisen Lernzielen, Fallbeispielen sowie Übungsaufgaben mit Beispiellösungen (vgl. Balzer/Beywl 2015). Als zweites didaktisches Prinzip wird der „Projektmethode“ folgend (vgl. Frey 2012, S. 218-234; Pfäffli 2015) ein – wenn möglich echtes – selbst oder z.B. von der Schulleitung in Auftrag gegebenes Evaluationsvorhaben geplant.

Während der zwei bis drei Präsenzphasen à 1 bis 1,5 Tage wird kaum noch Stoff vermittelt. Hingegen vertiefen und übertragen hier die Teilnehmenden das theoretisch Erarbeitete auf ihre eigene Bildungspraxis. Sie erstellen einen Evaluationsplan für ein konkretes Vorhaben in ihrem schulischen Kontext. Dabei arbeiten sie alleine, in Tandems oder in Kleinstgruppen und werden von den Kursleitenden beraten. Diese lenken den Lern- und Planungsprozess durch Arbeitsaufträge, die auf die Schritte des Evaluationsprozesses abgestimmt sind. Dabei auftauchende Fragen und Probleme können an den realen Beispielen der Teilnehmenden diskutiert werden. In selbstorganisierten Lerngruppen – hier agieren die Teilnehmenden präsent, per Telefon oder online – können darüber hinaus Fragen geklärt werden.

Die Präsenzphasen haben je ca. vierwöchigen Abstand zueinander. Damit bleibt dazwischen genügend Zeit für die Auseinandersetzung mit den neuen Lerninhalten bzw. das Lesen der entsprechenden Buchkapitel sowie für Vertiefungen und Anwendungen. Parallel beantworten die Kursleitenden Rückfragen via E-Mail oder Lernplattform.

Mit diesem Vorgehen wird die Kompetenzvermittlung direkt mit der praktischen Umsetzung verbunden. Es werden für die Schul- bzw. Unterrichtspraxis nützliche Evaluationen vorbereitet. Am Schluss des letzten Präsenztages verfügen alle Teilnehmenden (evtl. als Tandem) über einen umsetzbaren Evaluationsplan, eventuell bereits ergänzt um ein vorgetestetes Erhebungsinstrument. Dieser ist Grundlage für den zweiten Teil der Evaluationsweiterbildung, mit dem der Transfer in die Praxis sichergestellt wird.

Soll das empirische Methodenwissen ebenfalls erworben werden, sind zusätzliche, deutlich zeitintensivere Zusatzweiterbildungen einzuplanen. Oft wird es angeraten sein, dass sich Schulinterne etwa zur Konzeption eines Interviewleitfadens oder zur Konstruktion eines Fragebogens an eine evaluatorische (Methoden-)Beratung wenden, die über Kompetenzen des Typs III verfügt.

Im zweiten Teil folgt die Durchführung des geplanten Evaluationsprojektes in der Schule (teils beginnt diese schon im Verlauf des ersten Teils). Dabei ist auch zu klären, wieviel Wissen und Können insbesondere im empirisch-methodischen Bereich (Fragebogenerstellung/Interviewdurchführung) schulintern vorhanden ist, um die erforderlichen Datenerhebungsinstrumente zu erarbeiten und einzusetzen und um die resultierenden Daten auszuwerten. Je nach schulintern vorhandenen Kompetenzen und Ressourcen kann das Evaluationsprojekt mit mehr oder weniger intensiver externer Unterstützung durchgeführt werden.

An die (erstmalige) Durchführung kann sich eine Folge-Weiterbildung anschließen, in der die eigene Evaluationspraxis kritisch reflektiert und daran anschließend weitere oder vertiefte Evaluationskompetenzen aufgebaut werden.

Resümee

Dieser Beitrag skizziert einen Weg, wie das bislang enttäuschend wenig genutzte Potenzial der internen Schulevaluation für Bildung und Lernen entfaltet werden kann. Als zentrale Gelingensbedingung wird die Erweiterung der Evaluationskompetenzen von Lehrpersonen angesprochen. Über das Evaluations-*Wissen* und das handwerkliche Untersuchen-*Können* hinaus geht es um die *Geistes-Haltung*, dass systematisches Evaluieren zur pädagogischen Arbeit dazugehört. In praxisintegrierten Fortbildungen kann es gelingen, ein solches *Mindset* von Bildungsprofessionellen zu stärken, wenn der Nutzen des Evaluierens für die pädagogische Praxis nach kurzer Zeit eintritt und wenn er sich als nachhaltig erweist. Wie jede Bildungsanstrengung verlangt dies Investitionen, d.h., insbesondere Zeit für die eigene Fortbildung sowie Anreize im Berufsauftrag der Lehrkräfte.

Literatur und Internetquellen

- Altrichter, H./Posch, P. (2007): Lehrerinnen und Lehrer erforschen ihren Unterricht. Unterrichtsentwicklung und Unterrichtsevaluation durch Aktionsforschung. Bad Heilbrunn: Klinkhardt.
- Balzer, L./Beywl, W. (2015): evaluiert. Planungsbuch für Evaluationen im Bildungsbereich. Bern: hep.
- Bergmann, J./Sams, A. (2012): Flip Your Classroom: Reach Every Student in Every Class Every Day. Washington: International Society for Technology in Education.
- Bestvater, H./Beywl, W. (2015): Gelingensbedingungen der Selbstevaluation. In: Bolay, E./Iser, A./Weinhardt, M. (Hrsg.): Methodisch Handeln – Beiträge zu Maja Heiners Impulsen zur Professionalisierung der Sozialen Arbeit. Wiesbaden: VS, S. 133-144.
- Beywl, W. (2011): Modelle der Evaluation personenbezogener Dienstleistungen. In: Beck, I./Greving, H. (Hrsg.): Gemeindeorientierte Dienstleistungen. Behinderung, Bildung, Partizipation. Stuttgart: Kohlhammer, S. 169-173.

- Böttcher, W./Hense, J. (2016): Evaluation im Bildungswesen – eine nicht ganz erfolgreiche Erfolgsgeschichte. In: Die Deutsche Schule 108, H. 2, S. 117-135.
- Buhren, C.G. (2007): Selbstevaluation in Schule und Unterricht. Ein Leitfaden für Lehrkräfte und Schulleitungen. Köln: LinkLuchterhand.
- Burkard, C./Eikenbusch, G. (2000): Praxishandbuch Evaluation in der Schule. Berlin: Cornelsen Scriptor.
- Cousins, J.B./Goh, S.C./Elliott, C.J./Bourgeois, I. (2014): Framing the Capacity to Do and Use Evaluation. In: New Directions for Evaluation, H. 141, S. 7-23.
- Dahler-Larson, P. (2009): Learning-oriented Educational Evaluation in Contemporary Society. In: Ryan, K.E.C./Bradley, J. (Hrsg.): The SAGE International Handbook of Educational Evaluation. Thousand Oaks, CA: Sage, S. 307-322.
- Eckensberger, L.H. (2015): Integrating the Emic (Indigenous) with the Etic (Universal) – A Case of Squaring the Circle or for Adopting a Culture Inclusive Action Theory Perspective. In: Journal for the Theory of Social Behaviour 45, H. 1, S. 108-140.
- Frey, K. (2012): Die Projektmethode. Der Weg zum bildenden Tun. Weinheim: Beltz.
- Gärtner, H. (2015): Feedbackkultur auf mehreren Ebenen. In: Schulverwaltung spezial 17, H. 1, S. 34-35.
- Giel, S. (2016): Vom Nutzen der Programmtheorie in Evaluationen im Schulkontext. In: Die Deutsche Schule 108, H. 2. S. 149-162.
- Härri, R. (2015): Evidenzbasierte Unterrichts- und Schulentwicklung – Wenn Lehrpersonen zu Selbstevaluatoren werden und Unterrichtserfolge sichtbar machen. In: Erziehung und Unterricht 166, H. 1-2, S. 156-163.
- Hattie, J.A.C. (2014): Lernen sichtbar machen für Lehrpersonen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning for Teachers“. Besorgt von Wolfgang Beywl und Klaus Zierer. Baltmannsweiler: Schneider Verlag Hohengehren.
- Hattie, J.A.C. (2015): Lernen sichtbar machen. Überarbeitete und erweiterte deutschsprachige Ausgabe von „Visible Learning“. Besorgt von Wolfgang Beywl und Klaus Zierer. Baltmannsweiler: Schneider Verlag Hohengehren.
- Hattie, J.A.C./Masters, D./Birch, K. (2016): Visible Learning into Action. International Case Studies of Impact. London: Routledge.
- Huber, S.G./Wolfgramm, C./Kilic, S. (2013): Vorlieben und Belastungen im Schulleitungshandeln: Ausgewählte Ergebnisse aus der Schulleitungsstudie 2011/2012 in Deutschland, Österreich, Liechtenstein und der Schweiz. In: Jahrbuch Schulleitung, S. 259-271.
- IDES (Informationen Dokumentation Erziehung Schweiz) (2007): Evaluation der Bildungseinrichtungen und des Bildungssystems. Kapitel 9 des Schweizer Beitrags für die Datenbank „Eurybase – The Database on Education Systems in Europe“. URL: http://www.ides.ch/dyn/bin/12961-13438-1-eurydice_09d.pdf; Zugriffsdatum: 08.03.2016.
- King, J.A./Rohmer-Hirt, J.A. (2011): Internal Evaluation in American Public School Districts: The Importance of Externally Driven Accountability Mandates. In: New Directions for Evaluation, H. 132, S. 73-86.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2005): Bildungsstandards der Kultusministerkonferenz. München: Wolters Kluwer.
- Kühl, S. (2011): Organisationen – eine sehr kurze Einführung. Wiesbaden: VS.
- Lipowsky, F./Rzejak, D. (2015): Wenn Lehrer zu Lernern werden – Merkmale wirksamer Lehrerfortbildung. In: Lin-Klitzing, S./Di Fuccia, D./Stengl-Jörns, R. (Hrsg.): Auf die Lehrperson kommt es an? Beiträge zur Lehrerbildung nach John Hatties „Visible Learning“. Bad Heilbrunn: Klinkhardt, S. 144-160.

- Nevo, D. (2009): Accountability and Capacity Building: Can they Live Together? In: Ryan, K.E./Cousins, J.B. (Hrsg.): The SAGE International Handbook of Educational Evaluation. Los Angeles, CA: Sage, S. 291-303.
- Pfäffli, B.K. (2015): Lehren an Hochschulen. Eine Hochschuldidaktik für den Aufbau von Wissen und Kompetenzen. Bern: Haupt.
- Pietsch, M. (2011): Die Evaluationspraxis an deutschen Schulen. Ein Überblick aus empirischer Sicht. In: Schulmanagement 42, H. 4, S. 12-14.
- Russ-Eft, D.F./Bober, M.J./de la Teja, I./Foxon, M./Koszalka, T.A. (2008): Evaluator Competencies: Standards for the Practice of Evaluation in Organizations. San Francisco, CA: Jossey-Bass.
- Schön, D.A. (1983): The Reflective Practitioner. How Professionals Think in Action. New York: Basic Books.
- Scriven, M. (1972): Die Methodologie der Evaluation. In: Wulf, C. (Hrsg.): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München: Piper, S. 60-91 (engl. Originalausgabe 1966: The Methodology of Evaluation. Lafayette, IN: Purdue University).
- von Saldern, M. (Hrsg.) (2010): Selbstevaluation von Schule: Hintergrund – Durchführung – Kritik. Nordersted: BoD.
- Zenkel, S. (2015): Selbstevaluation und neue Autonomie der Schule. Kritische Anmerkungen zu vermeintlichen Selbstverständlichkeiten. Berlin: LIT.

Wolfgang Beywl, Prof. Dr., geb. 1954, Leiter der Professur für Bildungsmanagement sowie Schul- und Personalentwicklung an der Pädagogischen Hochschule Nordwestschweiz; wissenschaftlicher Leiter von Univation – Institut für Evaluation, Köln.
Anschrift: PH FHNW, Bahnhofstrasse 6, 5210 Windisch, Schweiz
E-Mail: wolfgang.beywl@fhnw.ch

Lars Balzer, Dr., geb. 1972, Leiter der Fachstelle Evaluation am Eidgenössischen Hochschulinstitut für Berufsbildung.
Anschrift: EHB IFFP IUFFP, Kirchlindachstrasse 79, 3052 Zollikofen, Schweiz
E-Mail: lars.balzer@ehb-schweiz.ch