

Hans Merkens

Evaluation in Schulen – Notwendigkeit und Grenzen

Zusammenfassung

Schulen werden gegenwärtig in vielen Bundesländern extern evaluiert und sind zusätzlich gefordert, sich intern zu evaluieren. Dazu wird ihnen Unterstützung von Dritten angeboten. Im Beitrag werden Formen der Evaluation und eine Abgrenzung zum Monitoring sowie das Verhältnis zur Schulinspektion dargestellt. Die mögliche Funktion von Evaluation im Rahmen der Neuen Steuerung wird erläutert. Nachdem Standards für die Evaluation aus zwei verschiedenen Perspektiven benannt worden sind, werden abschließend einige Beispiele zur Praxis der Evaluation angeführt.

Schlüsselwörter: externe Evaluation, interne Evaluation, Schulinspektion, Monitoring, Neue Steuerung, Standards für Evaluation

Evaluation in Schools – Necessity and Limits

Summary

Presently, schools are evaluated in many States of the Federal Republic of Germany externally and are in addition demanded to evaluate themselves internally. Support of third parties is offered to them. In this contribution forms of evaluation and a demarcation to monitoring as well as the relation to school inspection are shown. The possible function of evaluation within the scope of governance is explained. After evaluation standards have been named from two different perspectives, finally some examples are given for the practice of evaluation.

Keywords: external evaluation, internal evaluation, school inspection, monitoring, governance, evaluation standards

1. Einleitung

Erträge von Bildungseinrichtungen werden in Deutschland seit den letzten Jahren verstärkt kontrolliert (vgl. Döbert/Dedering 2008). Das war international bereits lange der Fall. Untersuchungen zu *School Effectiveness* und *School Improvement* sowie ähnlichen Themenstellungen haben eine Tradition (vgl. Scheerens/Creemers 1989;

Sammons 1999; Mortimore 1998; Reynolds et al. 2014). Im Bildungssystem wurde in Deutschland im 21. Jahrhundert von der Input- zur Output-Steuerung gewechselt. Aus diesem Grund muss man sich vermehrt der Ergebnisse des Arbeitens im Bildungssystem und der einzelnen Schulen versichern. In einem allgemeinen Verständnis dient dem Evaluation (vgl. Abs/Klieme 2005, S. 45).

2. Begriffsbestimmung

Eine klassische Beschreibung findet sich bei Gronlund (1968, S. 1f.), der Evaluation über das methodische Vorgehen in vier Schritten erfasst hat. Danach muss erstens die Bestimmung der Ziele erfolgen, zweitens müssen diese anschließend operationalisiert werden, es schließt sich drittens die Auswahl bzw. Konstruktion der Messinstrumente an; diese werden eingesetzt und ausgewertet. Abschließend kann viertens festgestellt werden, wieweit die Schüler und Schülerinnen die intendierten Ziele erreicht haben. Heute sind systematische Definitionen üblich. Pragmatisch hat die Gesellschaft für Evaluation (DeGEval 2004, S. 14) bestimmt: „Evaluation ist die systematische Untersuchung des Nutzens oder Wertes eines Gegenstandes.“ Ähnlich hat Köller (2015, S. 330) Evaluation „als jegliche Art der zielgerichteten und zweckorientierten Festsetzung des Wertes einer Sache“ definiert. Kuper (2005, S. 7) hat dem noch den Anwendungsaspekt hinzugefügt und drei Dimensionen unterschieden: „Jede Evaluation beinhaltet an Erfahrung gebundene Aussagen über Tatsachen, normative Aussagen als Grundlage für Bewertungen und prognostische Aussagen in der Form von Entscheidungen über zukünftiges Handeln.“ Während Köller (vgl. 2015, S. 333) den Graben zwischen Aussagen über Tatsachen und normativen Aussagen zu schließen versucht, indem er einen theoretischen bzw. konzeptionellen Überbau für jede Evaluation als notwendig erachtet, hält Kuper (vgl. 2005) das Verhältnis zwischen den verschiedenen Aussagetypen für ungelöst. Beide Varianten lassen erkennen, dass mit Evaluation nicht der Anspruch verbunden wird, eine Tatsache als solche abzubilden, sondern dass die Tatsache aus einem bestimmten Blickwinkel erfasst werden soll. Dessen Auswahl bedarf der Begründung. Anschließend kann dann festgestellt werden, ob man mit dem Ergebnis zufrieden ist oder eine Verbesserung für erforderlich hält. Wie diese zu erreichen ist, kann in der Regel der Evaluation nicht entnommen werden.

3. Zur Historie von Evaluationen

Evaluationen haben in den USA eine lange Tradition. Dabei waren zunächst Messungen von Schülerleistungen üblich. Guba und Lincoln (1989) haben ein Generationenmodell mit vier Phasen dargestellt (vgl. Kuper 2005, S. 28ff.; von Kardorff 2010, S. 241f.). Danach wurde in der ersten Phase gemessen; es wurden beispiels-

weise Schülerleistungen erhoben. In der zweiten Phase, die sie mit Beschreibung gekennzeichnet haben, ging es im Wesentlichen um die Erfassung und Erfüllung pädagogischer Programme. Die dritte Phase war durch eine Verbindung mit sozial- und bildungspolitischen Reformprogrammen gekennzeichnet; das Ziel war Politikberatung. Die letzte Phase, der sie sich auch verbunden fühlen, wird von ihnen als responsiv gekennzeichnet; hier steht der Einbezug der Akteure im Feld, das evaluiert wird, im Mittelpunkt. Während anfangs bei der Datenerhebung ausschließlich quantitative Methoden eingesetzt worden sind, hat vor allem beim vierten Schritt der Einbezug qualitativer Methoden zugenommen. Diese Phasen verlaufen auch parallel zueinander weiter; es geht in erster Linie darum, wie sich die Perspektiven im Laufe der Zeit von Phase zu Phase erweitert haben. Deutlich wird allerdings, dass nicht von einem einheitlichen Verständnis von Evaluation ausgegangen werden kann.

3.1 Formative und summative Evaluation

Die Unterscheidung von formativer und summativer Evaluation geht auf Scriven (1967) zurück, der sie im Kontext der Curriculumentwicklung eingeführt hat. Es handelt sich nach der Einteilung von Guba und Lincoln (1989) um Phasen der Beschreibung: Formative Evaluation bezeichnete Evaluationen im Prozess der Curriculumentwicklung; summative Evaluation bezog sich auf die Evaluation des Curriculums als Produkt. Köller (2015) hat die formative Evaluation – auf den Unterricht bezogen – als Prozessevaluation bezeichnet. Mit diesem Hinweis wird ein Bezug zum Unterricht sichtbar, bei dem klassisch zwischen Unterrichtsplanung, -durchführung und -nachbereitung unterschieden werden kann (vgl. Merkens 2010, S. 69ff.). Alle drei Phasen können evaluiert werden. Die summative Evaluation zählt zur dritten Phase. Formativen Evaluationen kommt darüber hinaus im Prozess der Schulentwicklung eine große Bedeutung zu.

3.2 Interne versus externe Evaluation

Der im Zusammenhang mit der Evaluation von Schulen gerne gebrauchte Begriff Selbstevaluation wird von einigen Autoren abgelehnt, weil nicht eindeutig zu definieren sei, wer mit „selbst“ gemeint sein könnte (vgl. Abs/Klieme 2005, S. 48; Berke-meyer/Müller/van Holt 2016, S. 212). Im englischen Sprachraum ist der Terminus „self-evaluation“ geläufig (vgl. Chapman/Sammons 2013). Wenig umstritten ist die Unterscheidung zwischen interner und externer Evaluation. Bei einer Befragung von Schulleitungen hat sich in Österreich herausgestellt, dass die interne Evaluation als wichtiger betrachtet wird, als dies für Ergebnisse externer Evaluationen zutrifft (vgl. Altrichter/Kemethofer 2015, S. 297). Internen Evaluationen kommt z.B. im Rahmen der Schulentwicklung eine große Bedeutung zu. Das gilt sowohl für die Evaluation des IST-Standes am Beginn als auch für die Begleitung des Prozesses und auch für

die Auswertung am Ende des Prozesses. Sie weisen in diesem Format eine hohe Ähnlichkeit zur Survey-Feedback-Methode auf, die in der Organisationsentwicklung eine Tradition hat (vgl. French/Bell 1990). Externe Evaluation wird von wissenschaftlichen Instituten, privaten Unternehmen oder Experten und Expertinnen in Schulen (vgl. Stockmann 2010, S. 22) oder von einer Bildungsverwaltung durchgeführt. Durch sie soll geprüft werden, wieweit vorgegebene Ziele sowohl bei der Organisation als auch bei den Ergebnissen einer Schule erreicht werden. Sie kann ein Auslöser für Schulentwicklung sein (vgl. Husfeldt 2011; Dederig 2012; van Ackeren/Klemm/Kühn 2015). Ziel von interner und externer Evaluation ist das Gewinnen von Steuerungswissen. Allerdings ist die Unterscheidung zwischen interner und externer Evaluation nicht immer trennscharf, weil auch bei internen Evaluationen externe Berater und Beraterinnen hinzugezogen und Instrumente eingesetzt werden können, die extern entwickelt worden sind.

4. Abgrenzungen

Evaluation wird gegenwärtig in vielen Fällen synonym mit Monitoring bzw. Schulinspektion verwendet (vgl. Pietsch 2010). Mit Monitoring wird allgemein die systematische Erfassung, Beobachtung, Messung bzw. Überwachung eines Vorgangs oder eines Prozesses bezeichnet. Systemmonitoring hat gegenwärtig im Bildungssystem Konjunktur. Im Unterschied zur Evaluation entfällt der Anspruch zu bewerten; dieser ist nur implizit insoweit enthalten, als die Auswahl der zu beobachtenden Tatsachen eine Wertung nach Wichtigkeit enthält, wenn z.B. Schülerleistungen im Bildungssystem stichprobenartig erhoben werden, wie das bei den internationalen Vergleichsstudien PISA (vgl. Deutsches PISA-Konsortium 2001) oder IGLU (vgl. Bos et al. 2003) der Fall ist, oder wenn über das Bildungssystem berichtet wird (vgl. Autorengruppe Bildungsberichterstattung 2014). Schulinspektionen sind für die einzelne Schule eine Variante der externen Evaluation (vgl. Holtappels 2008). Sie werden auch von den Initiatoren so verstanden (vgl. z.B. Schulinspektion Berlin 2016).

5. Evaluation und Neue Steuerung

Evaluation kann im Rahmen der Neuen Steuerung eine zentrale Funktion zukommen. Das trifft insbesondere für den Fall der evidenzbasierten Steuerung zu (vgl. Altrichter/Kemethofer 2015) und hängt mit der Hinwendung zur Output-Steuerung (vgl. Fend 2011, S. 4) sowie der zunehmenden Gewährung von Autonomie für die einzelne Schule zusammen (vgl. Dubs 2011; Abs/Klieme 2005, S. 46). Mit Evaluationen soll dann auf der Makroebene des Bildungssystems, aber auch auf der Mikroebene der einzelnen Schule die Wirksamkeit von Maßnahmen nachgewiesen werden, indem die Steigerung der Lernergebnisse als Erfolgskriterium dient (vgl.

Fend 2011, S. 9). Der Schwerpunkt der bisherigen empirischen Forschung liegt dabei auf der Mikroebene der einzelnen Schule. Hier lassen sich Quantität und Qualität der Bildungsangebote am einfachsten überprüfen. Kyriakides und Creemers (2008) haben mit einem multidimensionalen Modell den Einfluss von Faktoren auf der Klassenebene auf Schülerleistungen in Mathematik, Griechisch sowie Religion bestätigt. Viel schwieriger ist es allerdings, die Wirksamkeit von Maßnahmen der Steuerung im Mehrebenensystem des Schulwesens nachzuweisen (vgl. Maag Merki 2016; Fend 2011). Neben Schwierigkeiten der praktischen Realisierung hat Fend (ebd., S. 12) vor allem auf die theoretische Frage hingewiesen, wie Mechanismen, die erwünschte Steigerungen der Lernergebnisse bewirken sollen, identifiziert werden können. Ein Blick auf den Stand der Governance-Forschung belegt, dass die Annahme einer einfachen Wirkungskette nicht tragfähig ist (vgl. Altrichter/Maag Merki 2016; Berkemeyer 2010). Davon zu unterscheiden ist die Steuerung in der einzelnen Schule (vgl. Berkemeyer/Müller/van Holt 2016). Hier zeichnet sich ab, dass es vermehrt über die Nutzung von Feedback-Systemen zu Annäherungen an Evaluation kommen wird (vgl. Altrichter/Moosbrugger/Zuber 2016, S. 237).

6. Standards für die Evaluation

Dem Aspekt der Standards ist sowohl national als auch international Aufmerksamkeit geschenkt worden (vgl. z.B. Abs/Klieme 2005). Bei den Standards der Evaluation wird immer wieder auf die Arbeiten des Joint Committee (1994) verwiesen. Für Deutschland hat die DeGEval diese Standards übersetzt und leicht an die hiesigen Bedürfnisse angepasst (vgl. DeGEval 2004). Es handelt sich um ein additives Konzept, in dem aufgelistet wird, welche Standards einzuhalten sind und welche Kategorien in den einzelnen Dimensionen beachtet werden sollen. Für jede dieser Dimensionen werden Standards benannt. Die Reihenfolge der vier Dimensionen Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit lässt erkennen, dass der Anwendungsbezug bei diesen Standards dominiert. In der Tradition von Gronlund (1968) wird das methodische Vorgehen spezifiziert. Das steht in einem gewissen Spannungsverhältnis zu der von Sozialwissenschaftlern und Sozialwissenschaftlerinnen verfassten Literatur zur Evaluation bzw. deren Durchführung, bei der oft methodische und methodologische Aspekte obsiegen (vgl. Köller 2015). Vor diesem Hintergrund überrascht nicht, dass auch das Verständnis für Standards bei Evaluationen variiert. Abs und Klieme (vgl. 2005, S. 58) haben z.B. das Einhalten von Standards beim Einsatz von Leistungstests noch am ehesten als erfüllbar angesehen. Ihrer kritischen Bilanz ist aber zu entnehmen, dass allgemein das Einhalten von Standards bei der Evaluation von Schulen nicht erfüllbar zu sein scheint. Dabei interpretieren sie Standards aus der Sicht der Sozialwissenschaften. Das muss in einem gewissen Widerspruch zu den Erwartungen der Auftraggeber von Evaluationen einerseits und denen der Nutzer von Ergebnissen andererseits stehen.

Mit den Standards hängt die Auswahl der Methoden eng zusammen, die zur Gewinnung der Daten eingesetzt werden. Ursprünglich wurde die abhängige Variable als Schülerleistung mit quantitativen Methoden erhoben. Es wurden in der Regel Schulleistungstests eingesetzt. Diese Form der Messung ist auch heute noch üblich; allerdings hat sich eine wesentliche Änderung bei der Bezugsnorm ergeben. An die Stelle der normorientierten Messung ist eine kriteriale getreten. Es geht nicht mehr darum, eine relative Position in einer Normalverteilung zu bestimmen, sondern es wird angestrebt, das Erreichen bzw. den Grad des Erreichens von Lernzielen zu messen. Die angewendeten statistischen Methoden sind im Zuge dieser Veränderung immer elaborierter geworden (vgl. Köller 2015). Das weist den Nachteil auf, dass für Nichtsozialwissenschaftler und -sozialwissenschaftlerinnen die Nutzung der Daten erschwert wird. Generell kann davon ausgegangen werden, dass bei externen Evaluationen quantitative Methoden bevorzugt werden, weil auf diese Weise die Vergleichbarkeit der Ergebnisse von verschiedenen Untersuchungseinheiten leichter zu sichern ist.

Bei internen Evaluationen einzelner Schulen werden demgegenüber häufig qualitative Methoden eingesetzt. Das Spektrum reicht dabei von teilnehmender Beobachtung im Unterricht bzw. beim Schulleben über Interviews mit Personen – Schulleitung, Lehrkräften, Schülern und Schülerinnen, Eltern – und Gruppendiskussionen bis hin zu Dokumentenanalysen (Schulprogramm). Allerdings wird auch für diesen Typ von Evaluationen zunehmend ein Instrumentarium zur Verfügung gestellt, das eine Auswertung mit quantitativen Methoden ermöglicht. Das reicht von Fragebögen mit geschlossenen Antwortformaten bzw. wenigen offenen Fragen bis hin zu Anleitungen für die Kategorisierung von Beobachtungen im Schulleben bzw. Unterricht (vgl. Pietsch 2010). Während bei den quantitativen Methoden die Sicherung methodischer Standards im Zentrum steht, wird bei den qualitativen Methoden die Annäherung an die Besonderheiten des Untersuchungsobjektes angestrebt. In beiden Fällen ist die Nachvollziehbarkeit der Interpretationen eine wesentliche Bedingung. Dazu gibt es bei den qualitativen Methoden entsprechende Anleitungen, z.B. in Form der Auswertungsstrategie der qualitativen Inhaltsanalyse (vgl. Mayring 2010).

7. Zur Praxis der Evaluation

Abs und Klieme (vgl. 2005, S. 48ff.) unterscheiden bei der Evaluation zwischen einem Forschungs-, einem Entwicklungs- und einem Legitimations- sowie Kontrollparadigma. Während bei dem Forschungsparadigma die Idealform das Experiment darstellt (vgl. auch Köller 2015), wird bei dem Entwicklungsparadigma die Verbesserung der Institution angestrebt, und bei dem dritten Paradigma geht es im Kern um Rechtfertigung. Im Folgenden wird dem ersten dieser Paradigmen keine große Aufmerksamkeit geschenkt.

Gegenwärtig gibt es eine Flut von Evaluationen. Evaluation zählt in Schulen bereits zum Alltag: Sie ist in vielen Bundesländern in den Schulgesetzen vorgeschrieben und wird intern oder extern durchgeführt. Dabei sind häufig die Schule sowie die Organisation des Unterrichts von Interesse. Eine praxisorientierte Anleitung haben Chapman und Sammons (2013) verfasst. Sie enthält viele Anregungen, was warum zu veranlassen ist.

Zur Unterstützung der Schulen bei ihren internen Evaluationen ist in Deutschland ein Markt entstanden, in dem Externe ihre Unterstützung für die Durchführung von Evaluationen anbieten. Am sichtbarsten hatte sich die Bertelsmann Stiftung mit ihrem Instrument SEIS positioniert (vgl. Holland 2009). Dieses Instrument steht seit 2016 nicht mehr zur Verfügung. In vielen Fällen ziehen Schulen externe Berater und Beraterinnen hinzu.

Auf der Basis gesetzlicher Vorschriften müssen Schulen heute Schul- und Unterrichtsqualität kontrollieren. Maritz et al. (2006) haben ein Bewertungsbuch für Schulen auf der Basis der *European Foundation for Quality Management* (EFQM) vorgelegt, das den Vorteil bietet, verschiedene Facetten von Qualität zu berücksichtigen. Bei der praxisbezogenen Anleitung zu diesem Qualitätsmanagementsystem wird der Bezug zur Evaluation deutlich. Zur Evaluation werden z.B. in Berlin sowohl interne Evaluationen durchgeführt als auch Ergebnisse der externen Evaluation in der Form des Abschneidens bei Vergleichsarbeiten sowie der Auswertung der Berichte von Schulinspektionen herangezogen (vgl. Senatsverwaltung für Bildung, Jugend und Wissenschaft 2013).

Entgegen den bisherigen Erläuterungen zur Evaluation ist die Steuerung des Lernerfolgs bei den Schülern und Schülerinnen in vielen Fällen kein zentrales Thema. Oft werden vielmehr andere Aspekte des Schullebens evaluiert. Viele Evaluationen dienen der Zielsetzung, eventuell als notwendig angesehene Verbesserungen auf der Schulebene anzustoßen. Ko, Hallinger und Walker (2015) haben in Hongkong diese Zielsetzung als zu anspruchsvoll empfunden: Sie konnten erstens Effekte eher auf der Fachbereichs- als auf der Schulebene nachweisen und haben zweitens Daten aus Längsschnittuntersuchungen für notwendig erachtet, wenn Effekte bei den Schülerleistungen überprüft werden sollen. In der Regel stehen allerdings Daten nur aus Querschnittsuntersuchungen zur Verfügung.

Eine interessante Studie haben Scribner et al. (1999) vorgelegt, die am Beispiel der Einführung professioneller Lerngemeinschaften eine Evaluation dieses Prozesses beschrieben haben. Nach Park und Lee (2015) ist es allerdings im internationalen Vergleich schwierig, auf der Schulebene Bedingungen zu identifizieren, die das Entstehen professioneller Lerngemeinschaften unterstützen. In einer anderen Studie wird die Wirksamkeit eines Präventions-Curriculums zum Abbau aggressiven Verhaltens bei Schülern und Schülerinnen evaluiert (vgl. Schick/Cierpka 2013).

International ist ein häufiges Thema, welchen Beitrag Lehrkräfte am Erfolg von Schülern und Schülerinnen haben (vgl. Reynolds/Mujis/Trehanne 2003). Das ist bereits in vielen Studien untersucht worden (vgl. auch Hallinger/Heck/Murphy 2014). Hier war vor allem von Interesse, welchen Einfluss die Evaluation der Lehrkräfte auf den Erfolg der Schülerinnen und Schüler hat. Die Ergebnisse lassen Vorsicht angeraten sein. Das hängt sicher auch damit zusammen, dass die Modi der Evaluation kein einheitliches Bild abgeben (vgl. Murphy/Hallinger/Heck 2013). Diese Fragestellung ist in Deutschland bisher weniger untersucht worden. Ihr käme bei Fragen zur Personalentwicklung bzw. Weiterbildung Bedeutung zu.

Innerhalb der einzelnen Schule können somit bei der Evaluation die Schule insgesamt (Schulqualität, Schulinspektion), der Unterricht (Unterrichtsqualität) bzw. die Tätigkeit der Lehrpersonen im Fokus stehen. Falls Mängel entdeckt werden, sollen diese im Anschluss behoben werden. Eine Übersicht zum Forschungsstand bei internen Evaluationen findet sich bei Berkemeyer, Müller und van Holt (2016).

8. Zusammenfassung

Evaluation ist kein klar definierter Begriff. Es gibt vielmehr unterschiedliche Varianten und auch wissenschaftliche Traditionen, aus denen heraus ein Verständnis von Evaluation entwickelt worden ist. Deshalb gibt es bei der Praxis der Evaluation erhebliche Differenzen. Trotz dieser Einschränkung kommt Evaluationen innerhalb des Bildungssystems und bei der einzelnen Schule zunehmende Bedeutung zu. Weil inzwischen von der Kontrolle des Outputs bei Maßnahmen im Bildungssystem Hinweise auf mögliche Defizite erwartet werden, muss konsequenterweise evaluiert werden. Es wird allerdings darauf ankommen, das Verständnis in diesem Bereich weiterzuentwickeln. Die Selbstverständlichkeit, mit der der Begriff verwendet wird, trägt.

Literatur und Internetquellen

- Abs, H.J./Klieme, E. (2005): Standards für die schulbezogene Evaluation. In: Zeitschrift für Erziehungswissenschaft, 4. Beiheft, S. 45-62.
- Altrichter, H./Kemethofer, D. (2015): Neue Ansätze der Steuerung des Schulsystems und die Einstellung von Schulleitungen. In: Bildung und Erziehung 68, H. 3, S. 291-310.
- Altrichter, H./Maag Merki, K. (Hrsg.) (2016): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS.
- Altrichter, H./Moosbrugger, R./Zuber, J. (2016): Schul- und Unterrichtsentwicklung durch Datenrückmeldung. In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS.
- Autorengruppe Bildungsberichterstattung (Hrsg.) (2014): Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen. Bielefeld: Bertelsmann.
- Berkemeyer, N. (2010): Die Steuerung des Schulsystems. Wiesbaden: VS.

- Berkemeyer, N./Müller, S./van Holt, N. (2016): Schulinterne Evaluation – nur ein Instrument zur Selbststeuerung von Schulen? In: Altrichter, H./Maag Merki, K. (Hrsg.) (2016): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS, S. 209-234.
- Bos, W. et al. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster u.a.: Waxmann.
- Chapman, C./Sammons, P. (2013): School Self-Evaluation for School Improvement. What Works and Why. CfBT Education Trust. Reading: CfBT.
- Dedering, K. (2012): Schulinspektion als wirksamer Weg der Systemsteuerung? In: Zeitschrift für Pädagogik 58, H. 1, S. 69-88.
- DeGEval – Gesellschaft für Evaluation (2004): Standards für Evaluation. Köln: DeGEval – Gesellschaft für Evaluation e.V.
- Deutsches PISA-Konsortium (Hrsg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.
- Döbert, H./Dedering, K. (2008): Externe Evaluation von Schulen in vergleichender Perspektive – eine Einführung. In: Dies. (Hrsg.): Externe Evaluation von Schulen. Münster u.a.: Waxmann, S. 11-22.
- Dubs, R. (2011): Die teilautonome Schule. Ein Beitrag zur Ausgestaltung aus politischer, rechtlicher und schulischer Sicht. Berlin: Sigma.
- Fend, H. (2011): Die Wirksamkeit der Neuen Steuerung – theoretische und methodische Probleme ihrer Evaluation. In: Zeitschrift für Bildungsforschung 1, H. 1, S. 5-24.
- French, W.L./Bell, C.H. Jr. (1990): Organisationsentwicklung. Bern: Haupt.
- Gronlund, N.E. (1968): Readings in Measurement and Evaluation. New York: The Macmillan Company.
- Guba, E.G./Lincoln, Y.S. (1989): Fourth Generation Evaluation. Thousand Oaks, CA: Sage.
- Hallinger, P./Heck, R.H./Murphy, J. (2014): Teacher Evaluation and School Improvement: Analysis of the Evidence. In: Educational Assessment, Evaluation and Accountability 25, H. 1, S. 5-28.
- Holland, D. (2009): Interne Evaluation der Georg-Förster-Gesamtschule Wörrstadt mit Hilfe von SEIS. In: Appel, S./Ludwig, H./Rother, U./Rutz, G. (Hrsg.): Jahrbuch Ganztagschule 2009. Leben, Leisten, Lernen. Bad Schwalbach: Wochenschau Verlag, S. 162-175.
- Holtappels, H.-G. (2008): Externe Evaluation durch Schulinspektion und zentrale Prüfungen – Eine Einführung. In: Böttcher, W./Bos, W./Döbert, H./Holtappels, H.-G. (Hrsg.): Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive. Münster u.a.: Waxmann, S. 219-222.
- Husfeldt, V. (2011): Wirkungen und Wirksamkeit der externen Schulevaluation. In: Zeitschrift für Erziehungswissenschaft 14, H. 2, S. 259-282.
- Joint Committee on Standards for Educational Evaluation (1994): The Program Evaluation Standards: How to Assess Evaluations of Educational Programs. Thousand Oaks, CA: Sage.
- Ko, J./Hallinger, P./Walker, A. (2015): Exploring Whole School versus Subject Department Improvement in Hong Kong Secondary Schools. In: School Effectiveness and School Improvement 26, H. 2, S. 215-239.
- Köller, O. (2015): Evaluation pädagogisch-psychologischer Maßnahmen. In: Wild, E./Möller, J. (Hrsg.): Pädagogische Psychologie. Berlin/Heidelberg: Springer, S. 329-341.
- Kuper, H. (2005): Evaluation im Bildungssystem. Stuttgart: Kohlhammer.
- Kyriakides, L./Creemers, B.P.M. (2008): Using a Multidimensional Approach to Measure the Impact of Classroom-Level-Factors upon Student Achievement: A Study Testing the Validity of the Dynamic Model. In: School Effectiveness and School Improvement 19, H. 2, S. 183-205.

- Maag Merki, K. (2016): Theoretische und empirische Analysen der Effektivität von Bildungsstandards, standardbezogenen Lernstandserhebungen und zentralen Abschlussprüfungen. In: Altrichter, H./Maag Merki, K. (Hrsg.): *Handbuch Neue Steuerung im Schulsystem*. Wiesbaden: VS, S. 145-170.
- Maritz, B. et al. (2006): *Bewertungsbuch für Schulen. Eine Anleitung zur Bewertung der Schulqualität auf der Grundlage des Modells der European Foundation for Quality Management (EFQM)*. Bern: hep.
- Mayring, P. (2010): Qualitative Inhaltsanalyse. In: Flick, U./von Kardorff, E./Steinke, I. (Hrsg.) (2010): *Qualitative Forschung. Ein Handbuch*. Reinbek: Rowohlt, S. 468-475.
- Merkens, H. (2010): *Unterricht. Eine Einführung*. Wiesbaden: VS.
- Mortimore, P. (1998): *The Road to Improvement. Reflections on School Effectiveness*. Lisse: Swets & Zeitlinger.
- Murphy, J./Hallinger, P./Heck, R.H. (2013): *Leading via Teacher Evaluation: The Case of the Missing Clothes?* In: *Educational Researcher* 42, H. 6, S. 349-354.
- Park, J.-H./Lee, J.Y. (2015): *School-Level Determinants of Teacher Collegial Interaction: Evidence from Lower Secondary Schools in England, Finland, South Korea, and the USA*. In: *Teaching and Teacher Education* 50, H. 1, S. 24-35.
- Pietsch, M. (2010): *Evaluation von Unterrichtsstandards*. In: *Zeitschrift für Erziehungswissenschaft* 13, H. 1, S. 121-148.
- Reynolds, D./Mujis, D./Treharne, D. (2003): *Teacher Evaluation and Teacher Effectiveness in the United Kingdom*. In: *Journal of Personnel Evaluation in Education* 17, H. 1, S. 83-100.
- Reynolds, D./Sammons, P./Fraine, B.D./Damme, J. van/Townsend, T./Teddle, C./Stringfield, S. (2014): *Educational Effectiveness Research (EER): A State-of-the-Art Review*. In: *School Effectiveness and School Improvement* 25, H. 2, S. 197-230.
- Sammons, P. (1999): *School Effectiveness. Coming of Age in the Twenty-First Century*. Lisse: Swets & Zeitlinger.
- Scheerens, J./Creemers, B.P.M. (1989): *Conceptualizing School Effectiveness*. In: Dies. (Hrsg.): *Developments in School Effectiveness Research. Special Issue of the International Journal of Educational Research* 13, S. 691-706.
- Schick, A./Cierpka, M. (2013): *International Evaluation Studies of Second Step, a Primary Prevention Programme: A Review*. In: *Emotional and Behavioral Difficulties* 18, H. 3, S. 241-247.
- Schulinspektion Berlin (2016). URL: https://www.berlin.de/.../schulinspektion/handbuch_schulinspektion.pdf; Zugriffsdatum: 18.03.2016.
- Scribner, J.P./Cockrell, K.S./Cockrell, D.H./Valentine, J.V. (1999): *Creating Professional Communities in Schools through Organizational Learning: An Evaluation of a School Improvement Process*. In: *Educational Administrative Quarterly* 25, H. 1, S. 130-160.
- Scriven, M. (1967): *The Methodology of Evaluation*. In: Tyler, R./Gagne, R./Scriven, M. (Hrsg.): *Perspectives of Curriculum Evaluation*. Chicago, IL: Rand McNally & Co., S. 39-83.
- Senatsverwaltung für Bildung, Jugend und Wissenschaft (Hrsg.) (2013): *Handlungsrahmen Schulqualität in Berlin. Qualitätsbereiche und Qualitätsmerkmale*. URL: www.berlin.de/sen/bildung; Zugriffsdatum: 18.03.2016.
- Stockmann, R. (2010): *Rolle der Evaluation in der Gesellschaft*. In: Stockmann, R./Meyer, W. (Hrsg.): *Evaluation. Eine Einführung*. Opladen: Budrich, S. 25-53.
- van Ackeren, I./Klemm, K./Kühn, S.M. (2015): *Entstehung, Struktur und Steuerung des deutschen Schulsystems. Eine Einführung*. Wiesbaden: Springer VS.
- von Kardorff, E. (2010): *Qualitative Evaluationsforschung*. In: Flick, U./von Kardorff, E./Steinke, I. (Hrsg.) (2010): *Qualitative Forschung. Ein Handbuch*. Reinbek: Rowohlt, S. 238-250.

Hans Merkens, Prof. Dr. Dr. h.c., geb. 1937, emeritierter Universitätsprofessor für Erziehungswissenschaft (Arbeitsbereich Empirische Erziehungswissenschaft) an der Freien Universität Berlin.

Anschrift: FU Berlin, Fachbereich Erziehungswissenschaft und Psychologie, Fabeckstr. 69, 14195 Berlin

E-Mail: merken@zedat.fu-berlin.de

UNSERE BUCHEMPFEHLUNG



2014, 232 Seiten, geb., 39,90 €,
ISBN 978-3-8309-3146-1

E-Book: 35,99 €,
ISBN 978-3-8309-8146-6



Direkt zum
Buch

Michael Pfeifer (Hrsg.)

Schulqualität und Schulentwicklung

Theorien, Analysen und Potenziale

In diesem Band werden aktuelle Fragestellungen aus den Bereichen Bildungsmanagement und Schulentwicklung diskutiert. Einleitend werden theoretische Grundlagen und Forschungsbefunde zur Schulentwicklung betrachtet. Weiterhin werden Analysen zur Unterrichts- und Schulqualität in den Fokus genommen, zudem wird die Entwicklung von Ganztagschulen in Deutschland thematisiert. Abschließend werden relevante Entwicklungen zur Evaluation und zum Qualitätsmanagement im schulischen Kontext aufgezeigt.

Mit Beiträgen von Herbert Altrichter, Wolfgang Böttcher, Wilfried Bos, Magdalena Buddeberg, Christoph Burkard, Bert Creemers, Melanie Ehren, Bea Harazd, Sabine Hornberg, Marianne Horstkemper, Kevin Isaac, Leonidas Kyriakides, Stephan Maykus, Gerry McNamara, Joe O'Hara, Michael Pfeifer, Hermann Pfeiffer, Ernst Rösner, Hans-Günter Rolff, Wolfram Rollett, Klaus-Jürgen Tillmann, Stefanie van Ophuysen, Heike Wendt, Lothar Wigger und Ariane S. Willems.



www.waxmann.com