

Wolfgang Beywl

## **Evaluationskultur im Bildungsbereich**

### **Anregungen aus Organisationstheorie, Evaluationswissenschaft und Evaluationspraxis**

---

Bildungseinrichtungen sind durch begrenzte Steuerungsfähigkeit und Wertkonflikte gekennzeichnet. Sie verfügen neben ihrer Schau- und ihrer formalen über eine informale Seite. Evaluationskultur basiert auf evaluativem Denken. Dieses prägt die Organisationskultur bis ins Informale hinein. Der Autor führt ein Gedankenexperiment durch: Er überprüft die Geltung der Prinzipien eines evaluativen Denkstils in anders akzentuierten organisationalen Kulturen – denen der pädagogischen Reflexion und des Qualitätsmanagements. Abschliessend zeigt er an drei Bildungsorganisationen, woran sich Schritte hin zu einer Evaluationskultur zeigen können.

---

Categoria di articoli : Contributi scientifici

Citazione: Wolfgang Beywl, Evaluationskultur im Bildungsbereich, in: LeGes 30 (2019) 1

## Inhaltsübersicht

1. Die Organisationsperspektive
2. Organisationskultur im Bildungsbereich
3. Evaluationskultur in Spannungsfeldern der Bildung
4. Evaluatives Denken als Fokus
5. Gedankenexperiment zur Annäherung am Evaluationskultur
6. Prinzipien von Evaluationskultur
7. Chancen für Evaluationskultur

[Rz 1] Programme, Projekte und Massnahmen von Bildung und Erziehung werden ganz überwiegend von Organisationen wie Schulen und Hochschulen getragen. Sie konzipieren ihre Bildungsangebote, setzen sie um und evaluieren sie, womöglich unterstützt durch Organisationsexterne. Oder Dritte, z. B. Bund oder Kantone, geben Evaluationen in Auftrag, auch da sie den Löwenanteil der Bildungsfinanzierung aufbringen. Aufgrund der damit verbundenen öffentlichen Verantwortung ist Evaluation im Bildungsbereich stark verbreitet (Europäische Kommission 2015; MARBURGER/GRIESE/MÜLLER 2016).

[Rz 2] Evaluation beschreibt und bewertet für festgelegte Zwecke und nach begründeten Kriterien insbesondere Programme und Massnahmen. Dies geschieht systematisch, transparent sowie nachvollziehbar und basiert auf Daten und Informationen, die mit Hilfe sozialwissenschaftlicher Methoden gewonnen wurden (BALZER/BEYWL 2018, S. 25).

[Rz 3] Angesichts der starken Verbreitung von Evaluation im Bildungsbereich wird in diesem Beitrag die Ausprägung und Relevanz von Evaluationskultur in Bildungsorganisationen betrachtet. Der Beitrag versucht erste Antworten auf folgende Fragen zu geben: Was ist mit Evaluationskultur im Kontext von Bildungsorganisationen gemeint? Welche Rolle spielt darin das evaluative Denken? Wie positioniert sich eine dezidiert evaluativ geprägte Kultur gegenüber Kulturen der pädagogischen Reflexion und des Qualitätsmanagements? Welche Perspektiven gibt es für die Entwicklung von Evaluationskultur im Bildungsbereich?<sup>1</sup>

### 1. Die Organisationsperspektive

[Rz 4] Zahlenmässig und auch von der Relevanz für die Gesellschaft her dominieren im Bildungsbereich Organisationen, die öffentlich finanziert und gesteuert werden. Oft sind sie Bestandteile der öffentlichen Verwaltung oder unterliegen zumindest einer Regulierung durch diese. Beispiele sind insbesondere die Volks- und Kantonsschulen, die beruflichen Schulen, die Hochschulen und andere tertiäre Bildungseinrichtungen, Weiterbildungseinrichtungen sowie sonstige Träger von Bildungsmassnahmen. Deren Angebote und Programme werden oft mit Steuergeldern finanziert, manchmal auch über Stiftungen, die ebenfalls in einer gewissen öffentlichen Verantwortung stehen. Es gibt darüber hinaus auch im Markt tätige Bildungsorganisationen oder innerbetriebliche Bildungsabteilungen, für die die nachfolgenden Ausführungen auf ihre Übertragbarkeit zu prüfen wären.

---

<sup>1</sup> Christian Rüefli danke ich für sein konstruktiv-kritisches Gegenlesen von Vorfassungen dieses Beitrags. Er baut auf einem Vortrag auf, den der Autor auf Einladung von Emmanuel Sangra, Präsident der SEVAL, auf dem SEVAL-Kongress 2018 in Bern gehalten hat.

[Rz 5] Das Thema «Evaluationskultur» erhält besondere Relevanz, wenn man sich die informale Seite dieser Organisationen anschaut.

[Rz 6] Nach STEFAN KÜHL (2011) können drei Seiten einer Organisation unterschieden werden:

- die formale Seite mit ihren rechtlichen Grundlagen, etwa dem Arbeitsrecht, den Regelwerken, etwa zur Nutzung des Corporate Design, festgelegten Prozessen, einem Organigramm und der unvermeidbaren Hierarchie;
- die informale Seite mit den eingeschliffenen gegenseitigen Erwartungen der Organisationsmitglieder. Diese sind meist unausgesprochen und zeigen sich z. B.: im kurzen Dienstweg; im Austausch wichtiger Informationen im Treppenhaus; in ungeschriebenen, daher nicht einforderbaren Verhaltensrichtlinien; in Absprachen zwischen Mitgliedern zweier Organisationseinheiten, an denen die jeweiligen Vorgesetzten auch einmal nicht beteiligt sind;
- die Schauseite der Organisation, also das, was sie «mit Kalkül» insbesondere vom Formalen ihrer Umwelt präsentiert. Absicht ist, hohe Reputation in der Fachwelt zu erlangen, gegenüber Geldgebern und Behörden und anderen Stakeholdern Sicherheit und Verlässlichkeit zu signalisieren und – nicht zuletzt – potenzielle Kundinnen und Kunden und bzw. Nutzerinnen und Nutzer anzuziehen, also zum Beispiel Schülerinnen und Schüler und deren Eltern, künftige Studierende oder Weiterbildungsteilnehmende.

## 2. Organisationskultur im Bildungsbereich

[Rz 7] Mit Organisationskultur wird hier die Gesamtheit insbesondere des Informalen angesprochen, die das Handeln in Organisationen massgeblich beeinflusst. Dies kann in Übereinstimmung oder auch im Gegenstrom zum Formalen geschehen. Unter Organisationskultur werden üblicherweise verstanden: sichtbare Rituale und Mythen; gemeinsam geteilte Werte, Orientierungsmuster, sozial konstruierte Sinnsysteme; implizites Bewusstsein, oder unausgesprochene Grundannahmen, Einstellungen, Denkstile und Haltungen. Die Kultur entsteht wesentlich aus dem Interagieren, dem Reflektieren und dem Lernen in den kleinen Einheiten oder Teams einer Organisation (vgl. z. B. Schein 2010; SCHOLZ/HOFBAUER 1990; TREICHLER 1995). Organisationskultur ist wesentlich das Informale, das sich im alltäglichen Denken und Handeln realisiert.

[Rz 8] Wie kann dieses gestaltet werden? Manche Bildungspolitikern und -politiker, Leiterinnen und Leiter von Bildungsorganisationen, aber auch deren Beraterinnen und Berater haben einen Traum: die gezielte Formung der Organisationskultur und damit auch der informellen Beziehungsnetzwerke, der Netze von subtilen Anreizen und Sanktionen, Attraktionen und Aversionen bis hin zu den impliziten Handlungs- und Denkmustern der Organisationsmitglieder. Dies ist mit der Absicht verbunden, die Kultur so zu formen, dass sie die formale Seite optimal unterstützt und damit auch auf der Schauseite vorzeigbar ist.

[Rz 9] Doch wie man es auch angeht – ob durch noch stärkere Regulierung oder im Gegenteil durch weitgehenden Verzicht auf Regeln überhaupt –, das Informale sucht sich immer wieder Schleichwege und behauptet seinen Eigensinn. Abhängig von der Rolle in der Organisation oder von der Betrachtungsperspektive von aussen löst diese Beobachtung Sorge, Erleichterung oder Amüsement aus.

[Rz 10] In *Bildungsorganisationen* ist die Omnipräsenz des Informalen von besonderer Dynamik: Bildungs- und anspruchsvolle Lernprozesse können dann gelingen, wenn die beteiligten Personen – besonders die Lehrenden und die Lernenden – in einem verbindlichen Rahmen (dem

Formalen) flexibel, adaptiv, vielfach nicht vorhersehbar und überraschend kooperieren. Der – nicht selten übertriebene – pädagogische Konstruktivismus hat diese unverzichtbare Koproduktion als zentrale Gelingensbedingung für Bildung und Lernen theoretisch und empirisch aufgearbeitet (vgl. z. B. REICH 2008).

### 3. Evaluationskultur in Spannungsfeldern der Bildung

[Rz 11] Was ist angesichts dieser Ausgangslage in *Bildungsorganisationen* mit «Evaluationskultur» gemeint? Es ist eine Herausforderung, dies theoretisch und empirisch angemessen zu fassen. Die auf die Evaluation angewendeten Komplexitätstheorien (PATTON 2010; ACKERMANN 2016) verweisen auf zwei Besonderheiten von Bildungsorganisationen:

[Rz 12] Zum einen haben wir es im Falle der Bildung mit einem Technologiedefizit (LUHMANN/SCHORR 1982) zu tun. Die angestrebten Resultate von Bildungs- und Erziehungsprozessen können nicht gesichert «hergestellt» werden. Im Unterschied zu hoch komplizierten technischen Vorhaben, etwa dem Betreiben der ISS-Raumstation, sind menschliche Lern- und Sozialisationsprozesse nicht kumulativ optimierbar im Sinne einer (nahezu) garantierten Erreichung gesetzter Ziele.<sup>2</sup> Die Bildung jedes menschlichen Individuums beginnt wiederum bei null und ihr Verlauf ist weder vorhersehbar noch mit hinreichender Sicherheit steuerbar. Logische Modelle und Programmtheorien, wie sie inzwischen in viele Evaluationsansätzen ein obligatorisches Element sind, erfüllen hier wichtige Funktionen: Sie sollen das Bildungsprogramm und die gedachten Verbindungen zwischen Ausgangslage, Interventionen und gewünschten Resultaten explizieren, die Komplexität mässig reduzieren und Raum für das Erfassen von Unvorhergesehenem lassen, sowie Ankerpunkte für Fragestellungen und Datenerhebungen bieten, um schliesslich ein Lernen durch zyklisches «Testen» je weiter entwickelter Modelle zu ermöglichen (vgl. GIEL 2013).

[Rz 13] Zum zweiten haben wir es im Bildungsbereich mit einer bisweilen chronischen antinomischen Wertekonkurrenz (vgl. SCHLÖMERKEMPER 2017) zu tun. So ist es umstritten, welches die prioritären, im Zweifelsfall vorzuziehenden Ziele und damit auch Bewertungskriterien für gute Schulen, guten Unterricht, gute Bildung sind. Die Auseinandersetzungen um den Deutschschweizer Lehrplan 21, der Dauerclinch um Stabilisierung der Berufsbildungsquote versus Erhöhung der Studienanfängerquote, die Konkurrenz zwischen umfassender humanistischer Selbstbildung einerseits, Kompetenzerwerb für optimale Bewährung auf dem Arbeitsmarkt andererseits sind einige von ungezählten Beispielen. Diese Spannungen reichen bis auf die Mikroebene des Unterrichts (z. B. Gewähren von Belohnungen, wenn ja, wie) und führen bisweilen zur Austragung heisser Konflikte, vorzugsweise in der Lokalpresse, aber auch in den überregionalen Medien, zuletzt um den Königsweg zum Erwerb von (Recht-)Schreibkompetenzen in der Primarschule. Solche Wertkonflikte sind durch wissenschaftliche Forschung lediglich mässig milderbar. Sie stellen eine besondere Herausforderung für die Vermittlung einer Evaluationskultur auf der Schauseite von Bildungseinrichtungen und -systemen dar. In der Evaluation sind daher systematische

---

<sup>2</sup> In Feldern komplizierter technischer Systeme wie Kernkraftwerken, Flugzeugen oder dem Internet setzt man daher auf «Qualitätssicherung», gerahmt über nationale und internationale technische Normen, die beständig auf Basis umfassender Datenauswertungen verfeinert werden, und deren Einhaltung hohe Stabilität, minimale Fehleraten und geringste Risiken verspricht. Seit den 1980er-Jahren wird dieser Ansatz mit grosser Unbekümmertheit auf den Bildungsbereich übertragen (nur ein Beispiel: KÖLLER 2008). Dies, obwohl die Voraussetzungen für Qualitätssicherung in vorgenannten Sinne dort gar nicht gegeben sind (zur Kritik z. B. MÜNCH 2015, der den Begriff ebenfalls alltagssprachlich nutzt.).

Verfahren zur Bestimmung von Bewertungskriterien und zur Durchführung von Bewertungssynthesen von herausragender Bedeutung (vgl. MÄDER 2016).

[Rz 14] Evaluationskultur könnte als spezifische Ausprägung einer Organisationskultur bezeichnet werden, die über das Abstützen von Entscheidungen und organisationalem Lernen auf empirische Untersuchungen und Daten hinaus formale Verfahrensweisen bis hin zu informalen Veralltäglichungen etwa in Einstellungen und Haltungen aufweist, wie mit der begrenzten technologischen Beherrschbarkeit und den unvermeidbaren Wertkonflikten in Bezug auf Bildungsprozesse systematisch umgegangen wird.

#### 4. Evaluatives Denken als Fokus

[Rz 15] Ein Kernbestandteil einer solchen Evaluationskultur ist das evaluative Denken. Interne und externe professionell Evaluierende verfügen über Wissen, Können, Erfahrungen, Denkstile und Haltungen, die kennzeichnend sind für die Wissenschaft, die Kunst und das Handwerk der Evaluation.

*Evaluatives Denken ist am bedeutsamsten, wenn es in die Organisationskultur eingebettet ist. Das bedeutet, dass Menschen in der Organisation gegenseitig erwarten, dass sie sich gemeinsam einsetzen: für die Klärung von Schlüsselkonzepten, die Unterscheidung von Mitteln und Zielen, das Denken in Outcomes, die Prüfung der Qualität der verfügbaren empirischen Belege bezüglich Wirksamkeit, die Abstützung ihrer Standpunkte und Urteile auf empirische Belege.*<sup>3</sup>(PATTON 2014, Flashcard 1)

[Rz 16] In den Kompetenzprofilen, die von den Evaluationsgesellschaften verabschiedet (z. B. DeGEval 2004 und SEVAL 2012) oder in ihrem Umfeld entstanden sind, und auch in den Evaluationsstandards (z. B. DeGEval 2016 und SEVAL 2016) finden sich zahlreiche Hinweise darauf. Besonders einschlägig für Evaluationen im Umfeld von Bildungsorganisationen ist die Taxonomie von RUSSELL ET AL. (2008), die ein sehr breites, von professionellen Evaluierenden zu demonstrierendes Kompetenzbündel differenziert beschreibt und dessen Zusammensetzung begründet.

[Rz 17] Dass dabei der reflektierte Umgang sowohl mit dem Technologiedefizit als auch mit der Wertekonkurrenz zentral ist, wird aus dem einführenden Beitrag zum gleichnamigen Themenheft der *New Directions in Evaluation* von VO ET AL. (2018) zum evaluativen Denken deutlich:

- *Als eine wichtige Facette von Gesellschaft und Kultur besteht evaluatives Denken aus Interaktionen zwischen verschiedenen Wertsystemen.*
- *Als professioneller Wert in der Evaluation schliesst evaluatives Denken dieselbe Verpflichtung auf Haltungen und Verfahrensweisen ein, welche es unterstützen, dass Wahrheit im Kontext repräsentiert wird.*
- *Als Evaluationskompetenz spricht evaluatives Denken die Fähigkeit an, kreativ im Ungewissen, in Ambiguitäten und Komplexität zu navigieren.*

---

<sup>3</sup> Übersetzung des Autors. Originaltext: Evaluative thinking becomes most meaningful when it is embedded in an organization's culture. This means that people in the organization expect to engage with each other in clarifying key concepts, differentiating means and ends, thinking in terms of outcomes, examining the quality of evidence available about effectiveness, and supporting their opinions and judgments with evidence.

- *Als ein wichtiger Aspekt von Evaluationspraxis meint evaluatives Denken, dass man sowohl die Positionierung, die Annahmen, die Motivationen und die Verzerrungen von sich selbst als auch die von anderen in Betracht zieht.*<sup>4</sup>

## 5. Gedankenexperiment zur Annäherung am Evaluationskultur

[Rz 18] Um die Besonderheiten einer durch evaluatives Denken geprägten Evaluationskultur herauszuarbeiten wird ein Gedankenexperiment vorgeschlagen. Die in der nachfolgenden Tabelle aufgelisteten Prinzipien sollen das Nachdenken, also eine Art Meta-Reflexion, über Evaluationskultur und ihre Besonderheiten anregen. Dies soll gefördert werden durch das Kontrastieren mit zwei weiteren, im Bildungsbereich ebenfalls verbreiteten Ansätzen des Beschreibens und Bewertens, die zu anderen typischen Ausprägungen von Organisationskulturen führen:

- einer Kultur der pädagogischen Reflexion, des vertieften Nachdenkens über pädagogische Praxis. In solchen zuvörderst kognitiven Prozessen der (Um-)Strukturierung von Erfahrungen, Annahmen, Problemsichten oder Wissensbeständen spielen Werte eine wichtige Rolle. «Reflexion bzw. die Fertigkeit, Praxis zu reflektieren, wird als bedeutend angesehen, um professionelle Kompetenzen aufzubauen oder weiterzuentwickeln.» (AEPPLI/LÖTSCHER 2016, S. 79).
- einer Kultur des Qualitätsmanagements (QM), der präzisen Beschreibung von wünschbaren Eigenschaften von Prozessen der Leistungserstellung in Bildungsorganisationen, zusammen mit der Festlegung von Auditierungsverfahren<sup>5</sup>, mit denen die Implementation des QM-Systems in der Organisation überprüft wird. In den letzten 30 Jahren haben viele Bildungsorganisationen derartige QM-Systeme eingeführt, was sich auch daran zeigt, dass sie Qualitätslabel für die Aussendarstellung nutzen (vgl. KELLER et al. 2017).

[Rz 19] Ein weiteres hier lediglich angerissenes Thema, ist die Relationierung zu einer Forschungskultur: Forschung ist weitaus stärker als Evaluation auf Gewinn *neuen* Wissens ausgerichtet, mit dem Ziel, verallgemeinerbare d. h. Kontexte übergreifende Muster etwa der Verursachung (Attribution) zu entdecken bzw. zu überprüfen. Akademische Sozialforschung mit ihren eigenen Referenzsystemen (peer-reviewte Journale, hoch kompetitive Forschungsförderung) tendiert dazu, sich von den alltäglichen Informationsbedarfen der Praxis zu entkoppeln. Sie wird dann bedeutungslos für Lernen in einem bestimmten organisationalen Kontext, einem Kernanliegen der Evaluationskultur (vgl. OWEN 2003). Forschungs- und Evaluationskultur treffen an Hochschulen aufeinander, an denen Evaluation meist in einer Nische des formalen Qualitätsmanagements platziert ist. Dass man das Evaluative gerade hier mit Gewinn für die Organisation, ihre Mitglieder und weitere Stakeholder des Bildungsbereichs stärken könnte, ist an anderer Stelle ausgeführt (BEYWL et al. 2015).

---

<sup>4</sup> Übersetzung des Autors. Originaltext: As one facet of a society and culture, ET consists of interactions between multiple value systems; as a professional value in evaluation, ET entails the commitment to those virtues and procedures that support the representation of truth in context; as an evaluator competency, ET reflects the ability to creatively navigate uncertainty, ambiguity, and complexity; as one aspect of evaluation practice, ET involves the investigation of one's own as well as others positionality, assumptions, motivations, and biases.

<sup>5</sup> Vgl. kritisch zur Verwechslung von Auditierung und Evaluation allgemein POWER (1999); speziell für den Bereich Schule: HENSE und BÖTTCHER (2019).

## 6. Prinzipien von Evaluationskultur

[Rz 20] Zur exemplarischen Konkretisierung wünschenswerten evaluativen Denkens wird auf MICHAEL PATTONS Darstellung (2018a, S. 191 ff.) zurückgegriffen. Evaluationsfach- und Evaluationsprozessberatende (*evaluation facilitators*) sollen sich demgemäss an 13 Prinzipien<sup>6</sup> orientieren, wie sie in der nachfolgenden Tabelle 1 aufgelistet sind. Sie enthalten – selbst wieder evaluierbare – Aufforderungen an Evaluationsfachleute, die Evaluationsprozesse anleiten, beraten oder moderieren und somit die Evaluationskultur einer Organisation auf- und ausbauen helfen. Die Tabelle zeigt, inwieweit die 13 Prinzipien in den drei Zugängen Praxisreflexion (PR), Qualitätsmanagement (QM) und Evaluation (EV) erforderlich im Sinne von konstitutiv sind oder ob sie lediglich wünschenswert, also dispensierbar sind. Die entsprechende Einschätzung geht auf die vergleichende Beschäftigung des Autors mit der jeweiligen Standardliteratur zurück, wobei zur Reflexion erst in neuerer Zeit Überblicksartikel erscheinen.

[Rz 21] Eine Bemerkung zu nachfolgendem Vergleich vorab: Die Darstellung mag dazu verleiten zu resümieren, dass Evaluation überlegen im Sinne von umfassender ist als die beide anderen Zugänge. Das wäre ein voreiliger Schluss, denn es wird lediglich in eine Richtung geschaut – also in Bezug auf die Tabellenspalten von rechts nach links, aus der Perspektive der Evaluation. Als Fragestellung wird verfolgt: Inwiefern sind in beiden anderen Ansätze ebenfalls die Prinzipien der Evaluation verbindlich? Würde man die Fragerichtung umdrehen, also z. B.: «Inwiefern deckt Evaluation die Breite der Anforderungen des Qualitätsmanagements ab?», dann müsste man weitere Zeilen aus Sicht des Qualitätsmanagements (bzw. dann auch der Reflexion) ergänzen. Man würde womöglich bei der Evaluation Dispensierbares oder Lücken feststellen.

---

<sup>6</sup> Ihre Darstellung ist hier sehr kurz gehalten. Zu Begriff und Konzept von «Prinzip» im Zusammenhang mit Evaluation vgl. PATTON (2018b) und die Besprechung von BEYWL (2018). Verkürzt handelt es sich bei Prinzipien um Annahmen (nicht Gewissheiten), an denen sich Entscheidungs- oder Auswahl-situationen – hier: der Bildungsevaluation – orientieren. Sie halten formal (kurz, im Imperativ formuliert) und inhaltlich (ihs. empirische Überprüfbarkeit ihres Einhaltens) bestimmte Vorgaben ein.

Prinzip	PR	QM	EV
1. sei klar	+	+	+
2. sei intentional	+	+	+
3. sei sensitiv in Bezug auf den Kontext	+	+	+
4. mache Annahmen explizit	+	+	+
5. lege Rechenschaft ab	+	+	+
6. sei systematisch	o	+	+
7. schlussfolgere auf Basis empirischer Belege	o	+	+
8. handle in Bezug auf Interessenlagen und kulturelle Werte sensitiv	+	o	+
9. sei spezifisch, fokussiere und priorisiere	o	o	+
10. denke systemisch, versetze dich möglichst gleichberechtigt in viele Perspektiven	o	o	+
11. sei empfänglich für und adaptiere Komplexitätsüberlegungen, welche Planbares und Emergentes angemessen einbeziehen	o	o	+
12. kläre Kriterien und Kriterienpunkte mit Stakeholdern und expliziere sie (eher prinzipien- als regelbasiert)	o	o	+
13. sei wachsam in Bezug auf nicht-vorhergesehene Folgen	o	o	+

Legende: PR = Pädagogische Reflexion / QM = Qualitätsmanagement / EV = Evaluation + = erforderlich / o = ggfs. wünschenswert

**Tabelle 1: Geltung der Prinzipien evaluativen Denkens in QM und pädagogischer Reflexion**

[Rz 22] Gemäss der für das Gedankenexperiment gewählten Perspektive ergibt sich als Ergebnis einer groben Einschätzung Folgendes:

- Die ersten fünf Prinzipien (hellgrau eingefärbt) kennzeichnen die Basis eines expliziten und transparenten, zielgerichteten und begründeten Handlungsmodus, geleitet durch eine kritische und selbstkritische Grundhaltung, mit der Selbstverpflichtung auf Argumentation und Überprüfbarkeit. Sie sind allen drei betrachteten Zugängen gemeinsam.
- Die beiden Prinzipien 6 und 7 (hellgrün eingefärbt) zeigen, dass Qualitätsmanagement und Evaluation im Unterschied zu pädagogischer Reflexion auf geplante systematische Abläufe und empirische Daten verpflichtet sind. Evaluation geht dabei weniger schematisch vor. Z. B. können Erhebungsinstrumente noch während der laufenden Erhebung «lernen», wohingegen Qualitätsmanagement oft über lange Jahre an identischen Instrumenten festhält. Dies hängt auch damit zusammen, dass das Kriterium der Umsetzungstreue (*compliance*) von Massnahmen oder Interventionen im QM einen hohen Stellenwert hat. Auch hat dort der Vergleich von Ergebnissen über die Zeit oder zwischen Organisationen oder Organisationseinheiten, oft mit dem Zweck der Rechenschaftslegung, gegenüber der kurzfristigen Adaption von Erhebungsdesigns oder -methoden an sich verändernde Nutzungsabsichten der Stakeholder Vorrang. Überdies ist Qualitätsmanagement dominant auf die formale Seite bezogen, bedient die Schauseite der Organisation besonders intensiv. Auf die technologisch nie voll beherrschbaren und chronisch wertkonfliktären Kernbereiche von Bildung,

etwa den Unterricht, angewandt, wirkt sich Qualitätsmanagement sogar kontraproduktiv aus (vgl. BEYWL 2015).

Pädagogische Reflexion betont gegenüber strenger Systematik und datenbasierter Empirie das Situative, die Introspektion, das Individuelle und die womöglich auch affektiv eingefärbten Wechselbeziehungen in den Mikroprozessen des Lehrens und Lernens. Sie hat ihre Stärken auf der informalen Seite (vgl. AEPPLI/LÖTSCHER 2016). Sie basiert nur in Ausnahmefällen auf Daten und Dokumenten, und es gibt auch nur punktuell Berichte, an denen sich ein Reflexionsprozess oder -ergebnis durch Nichtbeteiligte nachvollziehen liesse.

Evaluation bewegt sich im Raum *zwischen* den beiden Polen – zwischen der fixierten, im Idealfall personenunabhängig einzuhaltenden möglichst vollständigen Prozessbeschreibung auf der einen Seite und dem stark personenabhängigen, tiefen und ausschnitthaften Reflektieren auf der anderen Seite. Eventuell kann sie, als Evaluationsfach- und Evaluationsprozessberatung auch zwischen beiden Polen vermitteln, eine gerade für den Bildungsbereich womöglich zentrale Funktion.

- Das Prinzip 8 (hellblau eingefärbt), welches auf kulturelle Sensitivität der Verfahrensverantwortlichen verweist, ist sowohl in der pädagogischen Reflexion<sup>7</sup> wie in der Evaluation<sup>8</sup> unverzichtbar.
- Die Gesamtheit der folgenden fünf Prinzipien 9–13 kennzeichnet die Besonderheit des evaluativen Denkens. Je nach Evaluationszweck und Nutzungsabsichten der Stakeholder werden aus den vielen möglichen Fragestellungen gezielt wenige relevante ausgewählt. Dies ist ein Unterschied zur offenen, situativ angelegten pädagogischen Reflexion einerseits und zum umfassend und breit sowie nach vorgegebenen Dimensionen «screenenden» (daher auch gelegentlich «total» genannten) Qualitätsmanagement andererseits. Evaluation strebt danach, viele wichtige Stakeholder-Gruppen in die Planung und deren laufende Anpassung einschliesslich der Festlegung von Fragestellungen und Bewertungskriterien einzu beziehen. So steigt die Chance, der systemischen Komplexität von Bildungsorganisationen und der Unvorhersehbarkeit von Entwicklungen gerecht zu werden. In der Evaluation sind systematische Verfahren des Bewertens entwickelt worden, von der Klärung der zugrundeliegenden sozialen Werte und der relevanten Kriterien bis hin zur Bewertungssynthese (vgl. DAVIDSON 2015; MÄDER 2016; BALZER/BEYWL 2018). Schliesslich bemühen sich Evaluierende, auch das zu erfassen, was ausserhalb des Expliziten und Formalen liegt. Dies sind insbesondere die unvorhergesehenen Folgen, seien diese schliesslich positiv oder negativ zu bewerten.

[Rz 23] Zusammenfassend für diesen allgemeinen Teil zur Evaluationskultur im Bereich von Bildungsorganisationen seien zwei Thesen formuliert:

- Der Zugang über die Bestimmungselemente des evaluativen Denkens ermöglicht es, konzeptionell zu klären, was Evaluationskultur ausmacht, wie die Evaluationskultur einer Bildungsorganisation gestaltet werden könnte.

---

<sup>7</sup> AEPPLI et al. (2016) nennen es «Auseinandersetzung mit anderen Perspektiven».

<sup>8</sup> Die Evaluationsstandards der Evaluationsgesellschaften weisen alle einen oder mehrere auf, in denen diese Perspektiven-, Werte- oder Interessenvielfalt betont wird. In der Version der SEVAL-Standards von 2016: «A10 Ethik: Sämtliche Tätigkeiten im Zusammenhang mit einer Evaluation erfolgen ethisch verantwortungsvoll und mit Sensibilität für die gesellschaftliche und kulturelle Vielfalt».

- Eine empirische Untersuchung der Evaluationskultur einer bestimmten Bildungsorganisation kann durch diese Bestimmungselemente orientiert werden, z. B. durch Erhebungen (Erfassung vorhandener Erzeugnisse und Verhaltensspuren, Beobachtung oder Befragung) zu Fragestellungen wie folgenden:
  - In welchem Ausmass demonstrieren organisationsinterne Evaluationsverantwortliche, dass sie ihre Beschreibungen und Bewertungen gemäss Prinzipien des evaluativen Denkens vornehmen?
  - Inwiefern führt dies in der jeweiligen Bildungsorganisation dazu, dass sich solche Denk- und Handlungsmuster unter deren Mitgliedern ausbreiten?
  - Wie selbstverständlich ist es für die Organisationsmitglieder, Annahmen zu explizieren, Bewertungskriterien zu klären und zu operationalisieren, Behauptungen empirisch zu belegen oder zu widerlegen, die Untersuchungen fortlaufend an die eintretenden Veränderungen anzupassen und datenbasierte Entscheidungen zu treffen?

[Rz 24] Entscheidend ist, dass solches Denken und Handeln nicht ausschliesslich im Formalen verankert ist, sondern auch im Alltäglichen, Beiläufigen, in den automatisierten Handlungsrou-tinen stattfindet – und dass eine Balance gehalten wird, namentlich gegenüber einem rigiden Einfordern solchen Denkens. Um dies empirisch zu erschliessen, wären ethnographische Studien im Alltag von internen Evaluationsstellen von Interesse; evtl. vergleichend angelegt mit der Fragestellung danach, welche Gemeinsamkeiten, welche Unterschiede zum Denken in Qualitätsmanagement-Einheiten oder pädagogischen Entwicklungsgruppen bestehen.

## 7. Chancen für Evaluationskultur

[Rz 25] Das Evaluative in den pädagogischen Alltag zu bringen, das «*Mainstreaming*» (SANDERS 2002), hat in den vergangenen zehn Jahren einen deutlichen Schub erhalten. Zu verdanken ist dies den Veröffentlichungen des neuseeländischen Bildungsforschers JOHN HATTIE. Seine Metana-lysen (vgl. HATTIE 2013 sowie BEYWL/ZIERER 2018) zu den Einflussfaktoren auf das Lernen ha-ben grosse Verbreitung gefunden. Das von ihm formulierte Prinzip, Lehrpersonen müssten zu Evaluierenden ihres eigenen Unterrichts werden (HATTIE 2014), hat der internen und der Selbst-evaluation sowie der datengestützten Schul- und Unterrichtsentwicklung besonders in Deutsch-land und der Schweiz starken Auftrieb gegeben.<sup>9</sup> Diese Evaluationsarten ersetzen zunehmend die externen Evaluationsverfahren, denen es – qualitätsmanagement- oder auditierungsnah – nicht gelungen ist, evaluatives Denken bis auf die Mikroebene des Unterrichts, des Lehrens und Lernens zu bringen. Wenn dies nun gelänge, dann könnte eine vitale Evaluationskultur im Bil-dungsbereich entstehen.

[Rz 26] Beispiele für das Mainstreamen von Evaluation und von evaluativem Denken gibt es inzwischen einige: Zum Abschluss seien drei skizziert:

- Die Schule Matzendorf im Kanton Solothurn führt – durch eine externe Evaluationsbera-terin begleitet – eine mehrperspektivische multimethodische interne Evaluation durch. Sie generiert Evaluationsprozessnutzen, indem sie fortlaufend empirisch basiert dazulernt. Sie

---

<sup>9</sup> Vgl. zu dieser für Fachkräfte im Bildungswesen wünschenswerten «evaluativen Grundhaltung» auch SCHÖBER et al. (2012) und die dort angegebene Literatur.

kombiniert ausgewählte pädagogische Prozesse mit Datenerhebungen. Sie macht die Ergebnisse gegenüber externen Stakeholdern wie Eltern und örtlichen Vereinen transparent (BEYWL/Joss 2017). Nachhaltigkeit wird dadurch begünstigt, dass der Kanton von der formalen Seite her in mehrjährigem Abstand eine interne Evaluation einfordert und ihre Umsetzung überprüft (nicht ihr Vorgehen oder ihre Ergebnisse). Dass sich Evaluationskultur nachhaltig auch im Alltag etabliert, dies hängt von einzelnen Personen ab, namentlich der Schulleitung.

- Die ebenfalls öffentliche Stonefields-Schule in Neuseeland führt, stark orientiert an JOHN HATTIES Plädoyer für formative Evaluation des Unterrichts, kontinuierlich Daten-Rückmeldezyklen durch. Diese stützen die Weiterentwicklung der Schulen, des Unterrichts bis hin zum hoch personalisierten Lernen der Schülerinnen und Schüler empirisch ab. Umgesetzt wird dies mit einem massgeschneiderten digitalen Rückgrat, auf der Basis von Vertrauen darein, dass die gewonnenen Daten strikt formativ genutzt werden – sie dienen ausschliesslich der Lernförderung, weder der Notengebung noch etwa der Begründung einer Nichtversetzung. In der Konsequenz erhält jede Schülerin, jeder Schüler morgens einen individuellen Stundenplan auf sein elektronisches Device. Aufgabenbearbeitungen werden hochladen, von den Lehrpersonen kommentiert und der so gemessene Lernfortschritt wird mit Hilfe der Software sichtbar gemacht. Nicht nur Lehrpersonen, sondern auch viele der teils sehr jungen Schülerinnen denken evaluativ, als Element der stark geförderten metakognitiven Strategien. Es gibt auch Ansätze, die Eltern in solche Zyklen einzubeziehen. Als eine Voraussetzung ist geschaffen, dass sie von zu Hause in die Lehr-Lern-Datenbanken der Schule hineinsehen können. Nachhaltigkeit einer schulischen Evaluationskultur nicht nur für diese Schule wird durch das konsequent formative Evaluationsdenken des nationalen Bildungssystems gefördert. Evaluatives Denken hat sich über die Schulleitung und die Initialgruppe hinaus stark ausgebreitet und ist für viele selbstverständlicher Bestandteil ihres alltäglichen Handelns.
- Das Institut Weiterbildung und Beratung (IWB) der Pädagogischen Hochschule FHNW mit ca. 100 Mitarbeitenden bildet Lehrpersonen weiter und berät Schulen. Es ist ihm innerhalb von ca. 5 Jahren gelungen, ein evaluativ begründetes und immer wieder angepasstes Kennzahlensystem zu erarbeiten. Hinzu kommen jährlich mehrere fokussierte internen Evaluationen von Programmen und Projekten, die von den je Verantwortlichen selbst durchgeführt werden. Evaluationsberichte werden zumindest intern zugänglich gemacht, teils auch öffentlich. Es ist die Ausnahme, dass grössere Programme ohne einen zumindest groben Evaluationsplan aufgesetzt werden. Der durch die Gesamtorganisation (Fachhochschule Nordwestschweiz) gegebene formale QM-Rahmen (gestützt auf das Modell der *European Foundation for Quality Management* – EFQM) wird in ein auf das Institut zugeschnittenes Qualitätsmanagement angepasst. Dieses weist viele evaluative Elemente auf. Evaluationskultur hat sich auch im Informalen verbreitet, doch hängt die Nachhaltigkeit stark von Personen ab, da das Hochschul-Dach vorrangig das QM stärkt.

[Rz 27] Die Beispiele für Ausprägungen von Evaluationskultur im Bildungsbereich sind punktuell und ihre Dauerhaftigkeit ist nicht gesichert. In Neuseeland ist diese jedoch wahrscheinlicher, denn der formale Rahmen, der sich in Bildungsgesetzen oder dem Nationalen Curriculum niedergelegt findet, begünstigt dauerhaft evaluatives Denken und Handeln, postuliert Prinzipien, sorgt für Instrumente, Ressourcen und Weiterbildung in Evaluation.

[Rz 28] Wünschenswert wäre, mehr Beispiele für Evaluationskulturen öffentlich zugänglich zu machen. Die Verbreitung von Evaluationskultur im Bildungsbereich ist auf Modelle angewiesen. Diese können von Evaluationsfach- und Evaluationsprozessberatenden für konkrete Bildungsorganisationen angepasst werden. Diese werden dann eher bereit sein, Evaluation in ihr selbstreferentielles System zu integrieren, wenn die Mitglieder überzeugt sind, dass sie mit vertretbarem Aufwand viel Nutzen für die Bildung und das Lernen der jeweiligen Zielgruppen schaffen kann. Evaluatives Denken wird sich dann nachhaltig etablieren können, wenn es selbstverständliches Element des beruflichen Alltags wird, sich also auch auf der informalen Seite der Organisation etabliert.

[Rz 29] Die neuen Lehrpläne mit ihrer Kompetenzorientierung und die Digitalisierung bieten Fenster der Gelegenheiten für mehr oder bessere Evaluationskultur im Bildungsbereich. Diese Gelegenheiten zu nutzen, erfordert den Einsatz von professionellen Evaluatoreninnen und Evaluatoren mit einem tiefen Verständnis von Lern- und Bildungsprozessen. Dabei kann – wie die Ausführungen zum evaluativen Denken zeigen – Evaluation im Umfeld von Organisationen als durch Bildung und Selbstbildung geprägter Prozess aufgefasst werden.

---

WOLFGANG BEYWL, Leiter der Professur für Bildungsmanagement sowie Schul- und Personalentwicklung an der Pädagogischen Hochschule FHNW und wissenschaftlicher Leiter Univation, Institut für Evaluation, Köln.

## Literaturverzeichnis

ACKERMANN, GÜNTER (2016): Evaluation und Komplexität. Wirkungskonstruktion in der Evaluation von Gesundheitsförderung und Prävention, Rheinfelden. [edco.unibas.ch].

AEPPLI, JÜRIG/LÖTSCHER, HANNI (2016): «EDAMA – Ein Rahmenmodell für Reflexion», in: *Beiträge zur Lehrerinnen- und Lehrerbildung*, Jg. 34, Heft 1, S. 78–97. [https://www.pedocs.de/volltexte/2017/13921/pdf/BZL\_2016\_1\_78\_97.pdf].

BALZER, LARS/BEYWL, WOLFGANG (2018): *evaluiert. – erweitertes Planungsbuch für Evaluationen im Bildungsbereich*, 2. überarb. Aufl. Bern: h.e.p. [http://www.evaluiert.lars-balzer.info/].

BEYWL, WOLFGANG (2015): «Schulinternes Qualitätsmanagement fördert den Unterricht am besten, wenn es sich selbst begrenzt», in: Katrin Kraus (Hg.), *Bildung von Lehrerinnen und Lehrern. Herausforderungen in Schule, Hochschule und Gesellschaft*, Opladen: Barbara Budrich, S. 109–128.

BEYWL, WOLFGANG (2018): «Patton, Michael Quinn: Principles-Focused Evaluation. The Guide», in: *Zeitschrift für Evaluation*, Jg. 17, Nr. 2, S. 393–397.

BEYWL, WOLFGANG/JOSS, ANNA-REGULA (2017): «Evaluieren mit Takt. Wie man Selbstevaluation in Schulentwicklung integrieren kann», in: *Pädagogik*, Jg. 69, Heft. 5, S. 28–32.

BEYWL, WOLFGANG et al. (2015): «Forschungsverständnis Pädagogischer Hochschulen – ein Diskussionsbeitrag», in: *Beiträge zur Lehrerinnen- und Lehrerbildung*, Jg. 33, Nr. 1, S. 134–151. [http://www.bzl-online.ch/archiv/heft/2015/1/134].

BEYWL, WOLFGANG/ZIERER, KLAUS (2018): «10 Jahre» Visible Learning «– 10 Jahre» Lernen sichtbar machen», in: *Pädagogik*, Jg. 70, Nr. 9, S. 36–41.

DAVIDSON, JANE E. (2005): *Evaluation methodology basics. The nuts and bolts of sound evaluation*, Thousand Oaks: Sage.

DeGEval, Deutsche Gesellschaft für Evaluation (2004): *Empfehlungen zur Aus- und Weiterbildung in Evaluation. Anforderungsprofile an Evaluatoreninnen und Evaluatoren, Alfter bei Bonn: DeGEval.*

DeGEval, Gesellschaft für Evaluation (2016): *Standards für Evaluation – Erste Revision auf Basis der Fassung 2002*, Mainz: DeGEval.

Europäische Kommission/EACEA/Eurydice (2015): *Qualitätssicherung im Bildungswesen: Strategien und Konzepte der Schulevaluation in Europa. Eurydice-Bericht*, Luxemburg: Amt für Veröffentlichungen der Europäischen Union.

GIEL, SUSANNE (2013): *Theoriebasierte Evaluation. Konzepte und methodische Umsetzungen*, Münster: Waxmann.

HATTIE, JOHN A. C. (2013): Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von «Visible Learning», besorgt von Wolfgang Beywl und Klaus Zierer, Baltmannsweiler: Schneider Hohengehren.

HATTIE, JOHN A. C. (2014): Lernen sichtbar machen für Lehrpersonen: Überarbeitete deutschsprachige Ausgabe von «Visible Learning for Teachers», Besorgt von Wolfgang Beywl und Klaus Zierer (englischsprachige Erstveröffentlichung 2012), Baltmannsweiler: Schneider Hohengehren.

HENSE, JAN ULRICH/BÖTTCHER, WOLFGANG (2019): «Evaluation – Ein starkes Instrument für die Entwicklung von Unterricht und Schule», in: Stephan Gerhard Huber (Hg.): *Jahrbuch Schulleitung*, Kronach: Carl Link, S. 247–260.

KELLER, HANS et al. (2012): Zur Einordnung von Q-Labels im Bildungsbereich. 2., aktualisierte Auflage, Bern: wbz cps.

KÖLLER, OLAF (2008): Bildungsstandards in Deutschland: Implikationen für die Qualitätssicherung und Unterrichtsqualität, in: Meinert A. Meyer/Manfred Prenzel/Stephanie Hellekamps (Hg.): *Perspektiven der Didaktik, Zeitschrift für Erziehungswissenschaft*, Sonderheft 9, Stuttgart: VS Verlag, S. 48–59.

KÜHL, STEFAN (2011): Organisationen – eine sehr kurze Einführung, Wiesbaden: VS Verlag für Sozialwissenschaften.

LUHMANN, NIKLAS/SCHORR, KARL-EBERHARDT (1982): «Das Technologiedefizit der Erziehung und die Pädagogik», in: *Zeitschrift für Pädagogik*, 25. Jg., Nr. 3, S. 345–365.

MARBURGER, HELGA (Hg.) (2016): Bildungs- und Bildungsorganisationsevaluation. Ein Lehrbuch. De Gruyter Oldenbourg: Berlin.

MÄDER, SUSANNE (2016): «Bewertungsverfahren in der Evaluation – Vorgehensweisen und Praxisbeispiele», in: Giel, Susanne et al. (Hg.): *Evaluationspraxis. Professionalisierung – Ansätze – Methoden*, Münster: Waxmann, S. 279–297.

MÜNCH, RICHARD (2015): Evaluation 2.0: von der wissenschaftlichen zur managerialen Qualitätssicherung?, in: *Bulletin der Schweizerischen Gesellschaft für Soziologie*, 147/148, S. 22–29.

OWEN, JOHN M. (2003): «Evaluation culture: A definition and analysis of its development in organizations», in: *Evaluation Journal of Australasia*, Jg. 3, No. 1, S. 43–47.

PATTON, MICHAEL Q. (2010): *Developmental evaluation: applying complexity concepts to enhance innovation and use*, New York: Guilford Press.

PATTON, MICHAEL Q. (2014): *Evaluation Flash Cards – Embedding evaluative thinking in organizational culture*, St. Paul, Minnesota: Otto Bremer Foundation.

PATTON, MICHAEL Q. (2018a): *Principles-focused evaluation: the guide*, New York: Guilford Press.

PATTON, MICHAEL Q. (2018b): *Facilitating evaluation. Principles in practice*, Los Angeles: Sage.

POWER, MICHAEL (1999): *The audit society. Rituals of verification*, Oxford: Oxford University Press.

REICH, KERSTEN (2008): *Konstruktivistische Didaktik: Lehr- und Studienbuch mit Methodenpool*, 4. Aufl, Weinheim: Beltz.

RUSS-EFT, DARLENE F. et al. (2008): *Evaluator competencies: standards for the practice of evaluation in organizations*, San Francisco: Jossey-Bass.

SANDERS, JAMES R. (2002): «On Mainstreaming Evaluation: Presidential Address», in: *American Journal of Evaluation*, Jg. 23, 3, S. 253–259.

SCHIEIN, EDGAR H. (2010): *Organisationskultur*, 3. Aufl., Bergisch Gladbach: EHP.

SCHLÖMERKEMPER, JÖRG (2017): *Pädagogische Prozesse in antinomischer Deutung. Begriffliche Klärungen und Entwürfe für Lernen und Lehren*, 1. Auflage, Weinheim: Beltz Juventa.

SCHÖBER, BARBARA et al. (2012): «Ergebnisorientierte Qualitätsentwicklung von Schule: Spezifische Kompetenzen von Lehrkräften, Schulleiterinnen und Schulleitern», in: Barbara Herzog-Punzenberger (Hg.): *Nationaler Bildungsbericht Österreich 2012*, Wien: Leykam, S. 111–142.

SCHOLZ, CHRISTIAN/HOFBAUER, WOLFGANG (1990): *Organisationskultur. Die vier Erfolgsprinzipien*, Wiesbaden: Gabler.

SEVAL, Schweizerische Evaluationsgesellschaft (2012): *Kompetenzanforderungen in der Evaluation. Arbeitsgruppe «Forschung und Lehre»*. Nach Konsultation der Mitglieder vom Vorstand der SEVAL am 28. August 2012 genehmigt, Bern: SEVAL.

SEVAL, Schweizerische Evaluationsgesellschaft (2016): *Evaluationsstandards der Schweizerischen Evaluationsgesellschaft*, Bern: SEVAL.

TREICHLER, CHRISTOPH (1995): *Kulturbewusste Unternehmensführung. Entwicklung eines Problemlösungskonzepts zur kulturbewussten Unternehmensführung unter besonderer Berücksichtigung der Instrumente der Erfassung, Beurteilung und Gestaltung der Unternehmenskultur*, Bern: Haupt.

VO, ANNE T./SCHREIBER, JACOB S./MARTIN, ASHLEY (2018): «Toward a Conceptual Understanding of Evaluative Thinking», in: *New Directions for Evaluation*, Jg. 2018, 158, S. 29–47. [<https://onlinelibrary.wiley.com/doi/abs/10.1002/ev.20324>].