

Patrizia Kühner

## Rezension zu

Behnke, K. & Steins, G. (2017). Möglichkeiten und Grenzen der Einflussnahme auf Unterricht durch die Qualitätsanalyse in Nordrhein-Westfalen: Perspektiven von Schulleitungen und Lehrkräften. *Empirische Pädagogik*, 31(2), 173–188.

## Kommentierter Kurzbefund

Durch externe Evaluationen alias „Schulinspektion“ oder „Qualitätsanalyse“ sollen Schulen angeregt werden, ihre Arbeit zu reflektieren und weiterzuentwickeln. In welchem Maße dies aus der Sicht von Schulleitungen und Lehrkräften in der Praxis gelingt, ist von großer Bedeutung, um Wirkungen dieser länderspezifisch gestalteten Formen von standardbasierter und evidenzorientierter Steuerung einschätzen zu können.

Behnke und Steins führten Interviews mit 30 Schulleitungen und befragten 387 Lehrkräfte von 17 Schulen in Nordrhein-Westfalen (NRW), indem diese vor und nach der Durchführung einer Qualitätsanalyse (QA) um ihre Einschätzungen zur Relevanz und Akzeptanz des Verfahrens gebeten wurden, das heißt, erhoben wurden Einstellungen zur QA.

Zentrale Einsichten sind, dass die QA für Schulleitungen bei der Argumentation und Legitimation von Entwicklungsmaßnahmen dienlich sein kann, allerdings benötigen sie zusätzliche Unterstützung für die Initiierung und Prozessgestaltung solcher Maßnahmen. Lehrpersonen problematisieren häufig die Methode der Unterrichtsbeobachtung, daher folgt der Ergebnisrezeption nicht zwingend Reflexion und Aktion. Dies wiederum plausibilisiert die mehrdeutige Forschungslage zu Wirkungen von externen Evaluationen, da die „selbstständige Schule“ eigenverantwortlich für den Prozess zwischen Bestandsaufnahme und Unterrichtsentwicklung zuständig ist. Zentrale Voraussetzungen für Qualitätsentwicklung auf Basis externer Evaluationen sind deshalb, die Schulleitungen und Lehrkräfte vom Verfahren zu überzeugen, Änderungsmotivation anzuregen und konkrete Unterstützung für den weiteren Prozess anzubieten.

Bedauerlicherweise fehlt es in vorliegender Studie an wesentlichen inhaltlichen Ausführungen, um die Ergebnisse und weiterführenden Gedanken besser nachvollziehen zu können. So fehlen beispielsweise Erläuterungen zur QA in NRW sowie Ausführungen zum theoretischen Konstrukt der Einstellung und zu den impliziten Annahmen, inwiefern sich Wirkungen von externen Evaluationen anhand der Einstellungen von Lehrkräften und Schulleitungen ermitteln lassen. Daneben sind die Ergebnisse wegen der lückenhaften Darstellung der Datengewinnung und -analyse, der unzulänglichen methodischen Reflexion der unterschiedlichen Herangehensweisen in den beiden Erhebungen sowie aufgrund der geringen Zahl an qualitativen Daten mit entsprechenden Einschränkungen zu rezipieren.

## Hintergrund

Einleitend referieren Behnke und Steins, dass zwischen 2004 und 2008 externe Schulbegutachtungen (sog. Schulinspektionen) in ganz Deutschland eingeführt wurden. Diese sind als Teil der Qualitätssicherung im Schulwesen zu verstehen. Analysen und Feedback der Inspizierenden dienen dabei als Grundlage für einen Reflexions- und Entwicklungsprozess schulischer Arbeit auf allen Ebenen, insbesondere aber hinsichtlich Unterrichtsverbesserung und -entwicklung.

Die Autorinnen sehen diese Entwicklung von Beginn an begleitet von kritischen Fragen, unter anderem danach, ob ministeriell initiierte Schulinspektionen das System Schule allgemein sowie Lehrpersonen im Besonderen zu einer (Unterrichts-)Reflexion und Unterrichtsweiterentwicklung anregen. Unterricht als „Kerngeschäft“ von Schule, das die meisten Akteurinnen und Akteure des Systems direkt berühre (Lehrpersonen, Schülerinnen und Schüler, Eltern), und Unterrichtsentwicklung seien von höchster Relevanz, da es um die Verbesserung einer Institution gehe, deren Ziel die Ausbildung zukünftiger Generationen von mündigen Bürgerinnen und Bürgern sei.

Die Autorinnen fassen zunächst den Stand der Forschung zu Schulinspektionen zusammen. Allgemein sei ein Zuwachs an Kontroll- und Steuerungsinstrumenten zu beobachten, die Standards sichern sollten. Diese Instrumente könnten unter anderem danach unterschieden werden, welche Konsequenzen sie nach sich zögen. In Europa obliege die Steuerung eher den Schulinspektionen, deren Ausgestaltung sich stark unterscheide und auf einem Kontinuum zwischen Unterstützung und Kontrolle angesiedelt sei. Die Bundesländer in Deutschland verfolgten dabei in erster Linie eine Entwicklung der Schulen „über Einsicht“.

### **Wirksamkeit von Schulinspektionen in Bezug auf Schul- und Unterrichtsentwicklung**

Die Forschungslage zur Wirksamkeit von Schulinspektionen hinsichtlich Schul- und Unterrichtsentwicklung ist gemäß den Ausführungen von Behnke und Steins nicht zufriedenstellend. So sei vor allem die Datenlage zur Wirkung auf die Schulentwicklung dünn, Ergebnisse ständen aus oder seien nicht klar in ihrer Aussage über die Wirkungsrichtung. Die Autorinnen halten zudem fest, dass es mindestens einer zweiten Evaluation bedarf, um Wirkungsmechanismen messen zu können. Allerdings verweise die vorhandene Datenlage durchaus auf mögliche positive Einflüsse von Schulinspektionen auf Schulentwicklung. Außerdem gelinge es nur einem kleinen Teil aller Schulen, sich ohne externe Rückmeldung selbst zu entwickeln.

Trotz der unklaren Forschungslage sprechen laut Behnke und Steins folgende drei Forschungsergebnisse für den Einsatz von externen Schulbegutachtungen: Das Feedback der Schulinspektion ...

1. sei ein entscheidender Faktor für die Bereitschaft zu einer positiven Verhaltensänderung;
2. werde von einer Mehrheit der Lehrkräfte als bedeutend für die Schulentwicklung angesehen;
3. unterstütze und bestärke Schulleitungen in ihrer Arbeit.

Hinsichtlich der Wirkung auf die Unterrichtsentwicklung konstatieren die Autorinnen einen stark divergierenden Forschungsstand und mehrere Desiderate. Als Erklärungsversuch führen sie an, dass durch die Schulinspektion zwar eine Evaluation stattfindet, aber im Nachgang kaum eine Beratung erfolgt und die weitere Unterrichtsentwicklung von den Schulen eigenverantwortlich zu gestalten ist. Eine Schule stehe hier also als Vermittler zwischen Ergebnissen der Inspektion und konkreter Unterrichtsgestaltung.

Abschließend geben Behnke und Steins einen Überblick zur Schulinspektion in Nordrhein-Westfalen

(NRW), die dort als „Qualitätsanalyse“ (QA) bezeichnet wird und seit einer Neuausrichtung im Jahr 2013 mehr partizipative Elemente enthält. Aus diesem Bundesland stammen auch die Daten für die vorliegende Studie. Die Schulinspektion in NRW ist in Bezug auf das oben beschriebene theoretische Kategorisierungssystem als unterstützendes Steuerungsinstrument mit niedrigen Konsequenzen anzusehen. Schneiden Schulen schlecht ab, findet eine zweite Begutachtung statt.

Die QA dient eher einer Bestandsaufnahme als einer Beratung und besteht im Wesentlichen aus den folgenden Elementen:

1. ein Abstimmungsgespräch vor der eigentlichen Evaluation, in dem schulische Gruppen ihre Angelegenheiten in das teils individuell erstellte Analysedesign einbringen können;
2. Besuchstage, bei denen vor allem Unterricht beobachtet wird;
3. Feedback zur durchschnittlichen Unterrichtsqualität durch mündliches Feedback an Schulleitung und Kollegium sowie durch einen schriftlichen QA-Bericht.

Die Ergebnisse sind nicht öffentlich. Auf ihrer Grundlage stimmt die Schulleitung mit der Schulaufsicht abschließend Zielvereinbarungen ab. Da die QA in dieser Form erst seit 2015 flächendeckend in NRW eingeführt ist und noch nicht alle Schulen evaluiert werden konnten, gibt es laut den Autorinnen noch keine repräsentativen Aussagen zur Langzeitwirkung.

Vor dem Hintergrund der unklaren Forschungslage zu den Wirkungen von externen Schulbegutachtungen auf die Schul- beziehungsweise Unterrichtsentwicklung sollen die Ergebnisse der nachfolgend erläuterten Erhebung Antworten auf folgende Fragen geben:

1. Welche Möglichkeiten und Grenzen gibt es bezüglich der Einflussnahme auf den Unterricht durch die QA in NRW?
2. Welche Problematiken und Verbesserungsbedarfe gibt es?

## **Design**

Es werden zwei Studien vorgestellt. Die Datengrundlage ergibt sich aus Erhebungen aus den Jahren 2013 bis 2016 an verschiedenen Schulen in NRW, die vom Ministerium für Schule und Weiterbildung beauftragt wurden.

### **Studie 1**

Für die erste Studie wurden 30 Schulleiterinnen und Schulleiter mit Hilfe von fokussierten Interviews zur Schulinspektion und Schulentwicklung befragt, wobei nicht alle Schulleitungen zum Zeitpunkt der Befragung bereits selbst eine QA erlebt hatten.

Fokussierte Interviews werden zur „Rekonstruktion subjektiver Bedeutungszuschreibungen beim Erleben spezifischer Situationen herangezogen [...]. Ausgangspunkt für ein fokussiertes Interview sind bestimmte Erfahrungen, Begegnungen oder Situationen, die eine Person erlebt und von denen der Forschende Kenntnis hat“ (Reinders, 2011, S. 89). Zunächst beobachten und dokumentieren Forschende die Situationen. Nach der Analyse dieser Dokumentation werden mit den Betroffenen Interviews geführt (sog. Prinzip der „retrospektiven Introspektion“).

Das Interview wurde mittels nicht weiter durch die Autorinnen erläuterten materieller und verbaler

Symbole gelenkt. Zur Nachvollziehbarkeit der beobachteten Reaktionen werden dann Beobachtung und Interview abgeglichen (vgl. Reinders, 2011, S. 90).

An der querschnittlichen Untersuchung nahmen Schulleitungen sowohl mit als auch ohne konkrete Erfahrungswerte zur QA an ihrer eigenen Schule teil. Sie leiteten gleichermaßen Berufskollegs, Förderschulen, Grundschulen, Gesamtschulen, Gymnasien, Hauptschulen und Realschulen. Durchschnittlich dauerte ein Interview etwa 43:30 Minuten.

Nach einer vollständigen Transkription wurde für die Analyse der qualitativen Daten die Software MAXQDA eingesetzt. Die Auswertungsmethode war die qualitative Inhaltsanalyse, in der durch Analyse des Materials Kategorien entwickelt und am Material überprüft werden (Gläser-Zikuda, 2011, S. 113). Im Fall der vorliegenden Studie wurde zuerst eine deduktive Kodierung auf der Grundlage der Fragebogenstruktur vorgenommen. Anschließend wurden aufgrund der in den Interviews benannten Themen Subkategorien identifiziert. Zur Prüfung der Kodierungen erfolgten eine Gegenkodierung und Validierung über eine nicht weiter benannte außenstehende Person sowie mit der Projektleitung.

## Studie 2

Das Thema der zweiten Studie war die Sichtweise von 387 Lehrenden nach dem Abstimmungsgespräch, aber noch vor den Besuchstagen. Diese wurde dann auch bei 326 Lehrenden zu einem zweiten Zeitpunkt nach der Schulinspektion gemessen. Die Erhebungsinstrumente bestanden aus einer Kombination von Fragebögen mit selbst erstellten Items, offenen Antworten und Interviews.

Die Stichprobe bestand aus einer Zufallsauswahl von Lehrpersonen, die an insgesamt 17 Grundschulen ( $n = 8$ ), Gymnasien ( $n = 6$ ) und Gesamtschulen ( $n = 3$ ) unterrichteten und sich in Regierungsbezirken befanden, die das neue QA-Verfahren bereits eingeführt hatten. Der quantitative Ansatz der zweiten Studie sollte ein grundsätzliches Stimmungsbild erheben. Dazu zählt die Akzeptanz der Lehrkräfte hinsichtlich der jeweiligen Phase der QA sowie die Relevanz der QA für die Lehrkräfte. Es wurden eine fünfstufige Likert-Skala (5 = *trifft völlig zu* bis 1 = *trifft gar nicht zu*) sowie die Option *ich weiß nicht* genutzt. Der qualitative Ansatz der Studie sollte Gründe für das Stimmungsbild eruieren.

## Ergebnisse

### Einstellungen von Schulleitungen zur QA (Studie 1)

Die Einstellungen der Schulleiterinnen und Schulleiter zur QA sind zu gerundet 23 % negativ, zu 27 % neutral und zu 50 % positiv. Drei Schulleitungen berichten eine positive Einstellungsveränderung infolge einer QA, eine Schulleitung eine negative. Die weiteren Einsichten aus den ermittelten Ergebniskategorien stellen die Autorinnen wie folgt zusammen:

- *Prozess*: Mehrere Schulleitungen vermissen neue Erkenntnisse durch das Inspektionsverfahren und wünschen eine dialogische Rückmeldung.
- *Personen*: Etwa ein Drittel der Schulleitungen äußert sich vorwiegend positiv zu den Qualitätsprüferinnen und Qualitätsprüfern, wobei ungefähr ebenso viele Kritik an deren Wahrnehmung einzelner Aspekte üben.
- *Instrument*: Etwa ein Drittel der Schulleitungen äußert Kritik und Zweifel an den Messmethoden, Einzelne betrachten die QA als Kontrollinstrument kritisch. Ein Drittel nimmt die QA positiv wahr, da sie Zielmarken setzt, ein Führungsinstrument darstellt und eine Argumentationsgrundlage gegenüber dem Kollegium ist. Etwa ein Sechstel vertritt eine sehr kritische und negative Haltung. Die Begründungsmuster werden in den Kategorien „hoher

Aufwand“, „Nachbereitung und Nachhaltigkeit“ sowie „belastende Aspekte“ zusammengefasst. Zusätzlich werden die Häufung ähnlicher Analyseergebnisse, ein fehlender Lerneffekt und das Ausbleiben von flächendeckenden Maßnahmen, die aus den Inspektionen bereits hätten resultieren können, aufgezählt.

- *Nutzung der Ergebnisse:* Ein Großteil der Schulleitungen weist auf die hohe Bedeutung eines kooperativen Evaluationsprozesses hin. Gewünscht werden unter anderem eine länger andauernde Begleitung bei der durch die Evaluation angeregten Schulentwicklung, ein regelmäßiger Austausch mit der Schulaufsicht, externe Evaluationen und mehr Fortbildungs- und konkrete Hilfsangebote. In diesem Sinne wird als mögliche Verbesserung vorgeschlagen, die QA als individuelles Beratungs- und Unterstützungsinstrument für die Steuerung von Entwicklungsprozessen aufzustellen.

In den Interviews wurden zudem Aussagen der 30 Schulleitungen zur Unterrichtsentwicklung (UE, 32 Aussagen) lokalisiert und in Unterkategorien eingeteilt, zum Beispiel „Bewertung von Unterricht durch die QA“ oder „UE als Aufgabe einer Schulleitung/Lehrkraft“.

Aus den in der Untersuchung dokumentierten Aussagen lässt sich einerseits ablesen, dass die QA Entwicklungsbedarfe beim Unterricht aufzeigen, Unterrichtsentwicklungsmaßnahmen anstoßen und legitimieren kann. Allerdings werden auch widrige Rahmenbedingungen, Widerstände von Lehrkräften gegenüber der QA und zu wenig passgenaue Unterstützung thematisiert sowie der Umstand, dass Unterrichtsentwicklung als Entwicklungspotenzial bei allen Schulen problematisiert wird und diesbezüglich Unterstützungsmaßnahmen fehlen.

### **Einstellungen von Lehrkräften zur QA (Studie 2)**

In der zweiten Studie wurden Einstellungen der Lehrkräfte zu zwei Zeitpunkten gemessen. Zum ersten Zeitpunkt, nach dem Abstimmungsgespräch, geben die Lehrpersonen im quantitativen Teil der Befragung mit einem Mittelwert von 2.04 ( $n = 333$ ,  $SD = 1.49$ ) an, dass die darin getroffenen Vereinbarungen kaum eine Auswirkung auf ihren eigenen Unterricht hätten. Anliegen aus dem Kollegium werden mit einem Mittelwert von 2.39 ( $n = 335$ ,  $SD = 1.68$ ) ebenfalls im Durchschnitt als eher weniger berücksichtigt ausgewiesen. Hinsichtlich der Abwägung von Aufwand und Nutzen sind 384 Lehrpersonen mit einem Mittelwert von 2.88 eher unentschieden.

Zu einem zweiten Erhebungszeitpunkt, nämlich nach den Besuchstagen der Inspektorinnen und Inspektoren, wurden die Lehrpersonen wiederholt befragt. Dazu wurden andere Items genutzt, von denen ausgewählte Ergebnisse zur Einstellung wie folgt ausgewiesen werden:

- Das Auftreten des Qualitätsteams wird als eher wertschätzend beurteilt ( $M = 3.92$ ;  $SD = 1.25$ ).
- Die Atmosphäre bei den Unterrichtsbesuchen wird als wenig unangenehm bewertet ( $M = 1.79$ ;  $SD = 1.37$ ).
- Aus Sicht der Lehrkräfte trifft es eher weniger zu, dass die Unterrichtsbeobachtungen die Realität abbilden ( $M = 2.61$ ;  $SD = 1.44$ ).
- Auswirkungen auf den Unterricht ( $M = 3.05$ ;  $SD = 1.27$ ) und auf die kollegiale Zusammenarbeit in den Fächern ( $M = 3.10$ ;  $SD = 1.44$ ) werden der QA nur teilweise zugesprochen.
- Die QA unterstützt die eigene Tätigkeit der Lehrkräfte eher nicht bis nur teilweise ( $M = 2.49$ ;  $SD = 1.21$ ).

Zusätzlich zu diesen Einschätzungen wurden zum zweiten Erhebungszeitpunkt auch Begründungen für die Einstellungen abgefragt. Von den 134 Angaben wurden für die vorliegenden Auswertungen 43

berücksichtigt, die sie sich auf Unterricht und Unterrichtsentwicklung (UE) bezogen. Auch diese Aussagen wurden in Kategorien zusammengefasst, zum Beispiel *Anzahl der Unterrichtsbesuche* (fünf Aussagen), *Realitätsnähe der Unterrichtsbeobachtung* (sieben Aussagen) und *Dauer und Zeitpunkt der Unterrichtsbeobachtung* (zehn Aussagen).

Die angeführten Beispielaussagen lassen erkennen, dass sich die Kritik der befragten Lehrkräfte unter anderem darauf richtet, dass bestimmte Unterrichtsphasen nicht gesehen wurden (etwa solche mit selbstgesteuertem Lernen), die Anzahl der Beobachtenden zu gering sei oder dass die konkrete Unterrichtsrealität von den Inspizierenden nicht angemessen gewürdigt wurde.

Bezüglich der Frage nach Möglichkeiten und Limitationen des Einflusses der QA auf den Unterricht fassen die Autorinnen folgende Einsichten der beiden Studien zusammen: Schulleiterinnen und Schulleiter fokussieren Unterrichtsentwicklung bereits vor der Schulinspektion und sehen in der Inspektion eine Unterstützung und Legitimation für Unterrichtsentwicklungsmaßnahmen.

Dagegen nehmen Lehrpersonen die QA als weniger anregend für die Unterrichtsentwicklung wahr. Sie nutzten die Befragung auch, um die Methoden der QA hinsichtlich Validität und Gestaltung fairer Rückmeldungen zu thematisieren. Dies steht in Verbindung mit der (niedrigen) wahrgenommenen Relevanz der QA. Die Autorinnen erwähnen an dieser Stelle ein Inspektionswirkungsmodell von Sommer (2011) mit den folgenden Phasen: Rezeption, Reflexion, Aktion und Evaluation. Dazu stellen sie die Hypothese auf, die Ergebnisnutzung der Rückmeldungen der QA sei möglicherweise bereits bei der Rezeption dieser Rückmeldungen stehen geblieben.

### **Implikationen für Schulleitungen**

Behnke und Steins leiten aus ihren Darstellungen folgende Implikationen für die Schulleitungen ab, damit eine QA wirken kann:

- Schulleitungen müssen sich ihrer bedeutenden Rolle bewusst sein, die sie für die Einstellung des Kollegiums zur QA spielen. Gruppen- und Motivationspsychologie-Schulungen sollten besucht werden, um Schul- und Unterrichtsentwicklungsprozesse gestalten, das heißt anstoßen und lenken zu können.
- Es müssen für Schulleitungen und Lehrpersonen bedarfsgerechte Fortbildungen sowie Unterstützungs- und Beratungsangebote im Anschluss an die QA bereitgestellt werden.
- Die Relevanzeinschätzung der QA und ihr wahrgenommener Nutzen von Seiten der Lehrkräfte könnten mit Maßnahmen, die sich positiv auf das primäre Tätigkeitsfeld der Lehrpersonen auswirken, gefördert werden.
- Es muss geklärt werden, wie Rückmeldungen möglichst wirksam – im Sinne von konkret und individualisiert für die einzelne Lehrperson sowie über die Inspektion hinaus – gestaltet werden können.

## **Diskussion und Einschätzung**

### **Hintergrund**

Die schlechten Ergebnisse der PISA-Studie aus dem Jahr 2001 waren Anlass für die Diskussion und Einführung neuer „Strategien, Verfahren und Instrumente der Steuerung von Schulen“ (Döbert, Rürup & Dederling, 2008, S. 63). Die Einführung von externen Evaluationen war Teil dieser Maßnahmen und ging mit der bereits in den 1990er-Jahren forcierten Outputsteuerung einher (vgl. Dederling, 2012, S. 76). Die

Frage nach der Wirkung von bildungspolitischen Maßnahmen ist entsprechend bedeutend und dabei gleichzeitig einer äußerst kritischen Rezeption ausgesetzt. Im Vordergrund von Schulinspektionen soll die Idee des Impulsgebens für schulische Qualitätsentwicklung stehen (vgl. ebd.).

Hervorzuheben ist, dass Behnke und Steins mit ihren deskriptiven Ergebnissen Erklärungsansätze anbieten, aus denen Thesen zu Wirkungsmechanismen hervorgehen können. In erster Linie bieten sie aber eine Beschreibung des Status quo.

Die Idee einer Untersuchung, wie die Einstellungen der betroffenen Akteurinnen und Akteure zu einer Qualitätsentwicklungsmaßnahme sind, ist aus dem Empfinden vieler an Qualitätsprozessen Beteiligter ein nachvollziehbarer kritischer Punkt für Entwicklungsprozesse. Der Lesende kann den Zusammenhang von Einstellung und Handlung und damit auch konkrete Ergebnisse in einem Entwicklungsprozess nachempfinden. Dass gerade mit Blick auf den Qualitätsbereich „Unterricht“ vor allem die Lehrpersonen eine Rolle spielen, liegt in dieser Betrachtungsweise auf der Hand und wird in der Untersuchung als Forschungslücke bisheriger systemischer Betrachtungen berücksichtigt.

Allerdings wird im Theorieteil nicht auf das psychologische Konstrukt „Einstellung“ eingegangen. Die impliziten Annahmen zum theoretischen Konstrukt der Einstellung in den Studien (z. B. der Zusammenhang von Einstellung und Verhalten; Haddock & Maio, 2014, S. 218ff.), die die wahrgenommene Wirkungsweise der QA in NRW erklären könnten, sollten für eine bessere Nachvollziehbarkeit und Einordnung der Ergebnisse expliziert werden.

## **Design**

Die Methoden passen grundsätzlich zum Erkenntnisinteresse (Interviews als offene Befragungsmethode zur Erhebung subjektiver Sichtweisen (vgl. Reinders, 2011, S. 86); computergestützte Analyse (z. B. Gläser-Zikuda, 2011, S. 117)). Auch die Kombination aus quantitativen und qualitativen Analysen für die zweite Studie ist durch die sich ergänzenden Einsichten passend begründet (Stichwort Methodentriangulation, also der Einsatz verschiedener Verfahren und deren Kombination für die Analyse eines Forschungsgegenstands, z. B. Gläser-Zikuda, 2011, S. 117).

Es fehlen allerdings in der Beschreibung der Instrumente, der Durchführung und der Ergebnisdarstellung wesentliche Informationen, was die Nachvollziehbarkeit einschränkt. Zudem fehlen Angaben, die die Güte der eingesetzten Verfahren beschreiben, wie zum Beispiel Angaben zur Übereinstimmung der Bewertenden (Intercoder-Reliabilität) oder zur Reliabilität der Skalen (Wie hoch ist die Messgenauigkeit?). Die Analyse erfolgt lediglich deskriptiv.

Hinsichtlich des psychologischen Konstrukts der Einstellung gibt es diverse Messansätze (vgl. z. B. Haddock & Maio, 2014, S. 212ff.). Eine Messung der Einstellung der Lehrpersonen mittels Frage-Items mit einer fünfstufigen Likert-Skala (z. B. Haddock & Maio, 2014, S. 213) im quantitativen Teil der zweiten Studie ist eine geeignete Methodenwahl (z. B. Gläser-Zikuda, 2011, S. 117). Mit Blick auf die theoretische Begründung des Designs und die Frage nach der Validität (Wird das gemessen, was gemessen werden soll?) wäre eine Erläuterung der Zusammenhänge der selbst formulierten Items mit dem Forschungsgegenstand, dem Einstellungs-Konstrukt, hilfreich.

Eine Aussage zur Repräsentativität der erhobenen Ergebnisse ist nicht leistbar, da bei der Beschreibung der Stichprobe für die erste Studie nicht angegeben wird, wie die Schulleitungen ausgewählt wurden, ob an deren Schule bereits eine QA stattgefunden hat und wie die Schulleitungen auf die Schulformen verteilt sind. Da nur Schulleitungen und Lehrkräfte aus NRW befragt wurden, können die Ergebnisse

auch nicht repräsentativ für die Schullandschaft in anderen Bundesländern sein.

Im Gegensatz zur ersten Studie handelt es sich bei der zweiten Studie um eine prinzipiell längsschnittliche Untersuchung. Bei der Betrachtung von Einstellungsentwicklungen könnte zum Beispiel ein Prä-Post-Design (Einstellungsmessung vor und nach der Intervention) Aufschluss über den Einfluss einer Intervention geben. Grundsätzlich wird der Zusammenhang der beiden Studien beziehungsweise die Gesamtanlage nicht erläutert. Somit bleibt unklar, warum Schulleitungen mit und ohne Erfahrung querschnittlich befragt wurden und die Lehrpersonen längsschnittlich. Genauso erklärungsbedürftig bleibt, warum sich die beiden Studien in Art und Anzahl der Schulformen unterscheiden und inwiefern die Einstellungen von Schulleitungen und Lehrpersonen deshalb gegenüberzustellen sind beziehungsweise zu gemeinsamen Aussagen zusammengefasst werden können.

## Ergebnisse

Die erste Frage (Welche Möglichkeiten und Grenzen gibt es bezüglich der Einflussnahme auf den Unterricht durch die QA in NRW?) wird mit Hilfe zweier Studien beantwortet. Die Ergebnisdarstellung zu Einstellungen der Schulleitungen hinsichtlich der QA ist zunächst eine Aufzählung von nicht weiter belegten quantitativen Einzeleinsichten aus den qualitativen Daten. Weiterführende Erläuterungen zum Zustandekommen dieser Zahlenwerte würden die Nachvollziehbarkeit der Ergebnisse unterstützen und auch die Frage klären, wie man aus einer querschnittlichen Untersuchung prozentuale Anteilsangaben zur Einstellungsentwicklung von Schulleitungen ableiten kann.

Sollten die Ergebnisse eine Unterscheidung von Schulleitungen mit und ohne QA-Erfahrung möglich machen, dann böte gerade diese Differenzierung interessante Einsichten hinsichtlich der Fragestellung der Untersuchung (z. B.: War die Einstellung mit oder ohne Erfahrung positiver oder negativer? Warum war das so? Was steuert die Einstellung?). Die Ergebnisdarstellung der Aussagen zu den Einstellungen der Schulleitungen ist nun wiederum nachvollziehbar und weist durchaus eine positive Wahrnehmung der QA auch zum Zweck der Unterrichtsentwicklung aus. Aufgrund fehlender Einsicht in Fragebogenstruktur und „der im Material benannten Themen“ sowie fehlender allgemeiner Erläuterungen der Kategorien bleiben die Ergebnisse der beiden Kodierungsschritte allerdings exemplarisch und teilweise schwer zu interpretieren.

Generell ist anzumerken, dass die Anzahl der Aussagen sehr gering ist: Bei 32 Aussagen von 30 Schulleitungen haben sich diese im Schnitt nur einmal zu dem Thema Unterrichtsentwicklung geäußert. Ohne Angaben dazu, inwiefern sich Schulleitungen mit und ohne Inspektionserfahrung explizit unterscheiden, ist die ausgewiesene Einsicht mit Fokus auf das Thema Unterrichtsentwicklung nicht nachvollziehbar.

Die abschließende Darstellung der Einstellungen der Schulleitungen bleibt rein deskriptiv. Auch wenn die Einstellung positiv ist, muss man nicht im Status quo verharren. Anstelle dessen könnte man fragen, was zu einer positiven Einstellung führt und wie man diese noch weiter begünstigen könnte.

Vergleichbar zur Kritik an der ersten Studie ist auch für die zweite Studie zu fragen, inwiefern es sich hier bei beiden Erhebungszeitpunkten um „auszugsweise Ergebnisse“ oder „ausgewählte Ergebnisse“ handelt beziehungsweise was die Auswahl der vorgestellten Ergebnisse geleitet hat. Selbst in der sich anschließenden Diskussion sprechen die Autorinnen global von „auszugsweise dargestellten Ergebnissen“, ohne die Auswahl zu erläutern oder daraus resultierende Limitationen kritisch zu betrachten. Beim zweiten Frage-Item der ersten Erhebung gibt es widersprüchliche Angaben der Standardabweichungen (vgl. Tabelle S. 181:  $SD = 1.68$ ; Text S. 180:  $SD = 2.38$ ).



Für die Ergebnisdarstellung der Einstellungen der Lehrpersonen zur QA und dem Unterricht ist eine vergleichbare Kritik wie bei der ersten Studie zu üben: Die tabellarische Ergebnisdarstellung ist nicht weiter oder nur mit Einzelaussagen exemplarisch erläutert und lässt so viel Interpretationsspielraum. Ebenso gibt es nur wenige Antworten (2–3 Aussagen). Generell sind die Ergebnisdarstellungen in Tabelle 2 bis 4 schwer nachzuvollziehen, wenn man nicht mehr Einblick darin hat, was genau bei den Abstimmungsgesprächen und bei den Unterrichtsbeobachtungen gemacht wurde. Mehr Hintergrundinformationen (z. B. in den theoretischen Ausführungen zur QA) wären hilfreich, um die Ergebnisse überhaupt einordnen zu können. Es bleibt letztlich unklar, ob die Ergebnisse der QA selbst so unkonkret sind, dass die eher globalen Einstellungen nicht konkreter herausgearbeitet werden konnten, oder ob das bereits an der wenig spezifisch erläuterten Art und Weise der Datenerhebung lag.

Schlussendlich bestätigen die Autorinnen mit den Daten ihre Eingangshypothese zur Erklärung des divergierenden Forschungsstandes: Unterrichtsentwicklung muss im Anschluss an eine Inspektion eigenverantwortlich von den Schulen gestaltet werden. Eine Schule steht also als Vermittler zwischen Ergebnissen der Inspektion und konkreter Unterrichtsgestaltung. Für diesen Prozess wünschen sich die Beteiligten mehr Unterstützung und Begleitung. Unberücksichtigt bei der Zusammenfassung bleiben differenzierte Implikationen, die sich gegebenenfalls aus dem Datenmaterial hätten ergeben können (z. B. unterschiedliche Implikationen für einzelne Schulformen). Der Wunsch nach einer differenzierten Theorieklärung wird auch beim Nachvollziehen der Implikationen erneut geweckt: Die hervorgehobene Bedeutung der Einstellungen der Lehrpersonen und der Einflussnahme der Schulleitung auf das Kollegium bezieht sich unmittelbar auf ein nicht ausformuliertes Verständnis von Konstrukten wie „Motivation“ oder „Einstellung“.

Ebenso wäre eine Anbindung der Forschungsergebnisse an den anfänglich ausgewiesenen Forschungsstand zu den Wirkfaktoren wünschenswert. Konkrete Vorschläge dazu, wie man positiven Einfluss auf die Einstellungen der Lehrpersonen nehmen könnte, bleiben aus. Dies wäre aber gerade vor folgendem denkbarem Einwand wünschenswert: Welches Instrument könnte denn besser zu konkreten und individuellen Veränderungen anregen als eine kriteriengeleitete Unterrichtsbeobachtung? Die Vorbehalte der Lehrpersonen gründen vermutlich auf der fragwürdigen Verwendung des Instruments der Unterrichtsbeobachtungen im Gesamtprozess der QA: Unterrichtsbeobachtungen werden genutzt, um daraus „Werte“ und darauf aufbauendes Feedback für die gesamte Schule zu generieren, die dann wiederum von der Schule interpretiert und zurückgespiegelt werden. Führt ein solcher „Umweg“ nach dem Stille-Post-Prinzip nicht zu mehr Missverständnissen und Frustration?

Eine ausgebliebene abschließende kritische Auseinandersetzung mit der Methode und der Methodologie hätte den ein oder anderen hier geäußerten Kritikpunkt gegebenenfalls aufklären können.

## **Reflexionsfragen für die Praxis**

Nachfolgende Reflexionsfragen sind ein Angebot, die Befunde der rezensierten Studie auf das eigene Handeln als Lehrkraft oder Schulleitungsmitglied zu beziehen und zu überlegen, inwiefern sich Anregungen für die eigene Handlungspraxis ergeben. Die Befunde der rezensierten Studien sind nicht immer generalisierbar, was z. B. in einer begrenzten Stichprobe begründet ist. Aber auch in diesen Fällen können die Ergebnisse interessante Hinweise liefern, um über die eigene pädagogische und schulentwicklerische Praxis zu reflektieren.

### Reflexionsfragen für Lehrkräfte:

- Inwiefern bestimmt meine Haltung gegenüber externer Evaluation den Grad meiner Unterrichtsinnovationen positiv oder negativ?
- Welche Kriterien für guten Unterricht werden von außen an Schulen herangetragen?
- Wer oder was motiviert mich, meinen Unterricht weiterzuentwickeln?
- Welche Art von Feedback, Hilfestellungen und Fortbildungsangeboten wünsche ich mir von wem, um meinen Unterricht und damit meine Schule gesamthaft weiterzubringen?

### Reflexionsfragen für Schulleitungen:

- Wie kommuniziere ich Evaluationsergebnisse (z. B. als daraus resultierende Entwicklungsstrategie) transparent und systematisch an die betroffenen Gruppen in meiner Schule?
- Wie kann über die gemeinsame Analyse und Reflexion von Evaluationsergebnissen eine systematische Schul- und Unterrichtsentwicklung angestoßen und begleitet werden?
- Inwiefern bestimmt meine Haltung gegenüber der Schulinspektion die Art und Weise der Kommunikation der Ergebnisse in „meinem“ System?
- Welche Unterstützungsangebote möchte ich bei der Umsetzung von Feedback aus Schulevaluationen nutzen? Welche Möglichkeiten bringe ich diesbezüglich selbst mit und wo sehe ich eigenen Fortbildungsbedarf?

## Literatur

Dedering, K. (2012). Schulinspektion als wirksamer Weg der Systemsteuerung? *Zeitschrift für Pädagogik*, 58(1), 69–88.

Döbert, H., Rürup, M. & Dederich, K. (2008). Externe Evaluation von Schulen in Deutschland – die Konzepte der Bundesländer, ihre Gemeinsamkeiten und Unterschiede. In H. Döbert & K. Dederich (Hrsg.), *Externe Evaluation von Schulen – Historische, rechtliche und vergleichende Aspekte* (S. 63–151). Münster: Waxmann.

Gläser-Zikuda, M. (2011). Qualitative Auswertungsverfahren. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Strukturen und Methoden* (S. 109–119). Wiesbaden: VS Verlag für Sozialwissenschaften.

Haddock, G. & Maio, G. R. (2014). Einstellungen. In K. Jonas, W. Stroebe & M. Hewstone (Hrsg.), *Sozialpsychologie* (S. 197–229). Berlin: Springer.

Reinders, H. (2011). Interview. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Strukturen und Methoden* (S. 85–97). Wiesbaden: VS Verlag für Sozialwissenschaften.

## Rezensent/-in

Dr. Patrizia Kühner, Dipl.-Hdl., Studienrätin an der Josef-Durler-Schule Rastatt.

## Zitiervorschlag

Kühner, P. (2021). Rezension zu Behnke, K. & Steins, G. (2017). Möglichkeiten und Grenzen der Einflussnahme auf Unterricht durch die Qualitätsanalyse in Nordrhein-Westfalen: Perspektiven von Schulleitungen und Lehrkräften. *Empirische Pädagogik*, 31(2), 173–188. *Forschungsmonitor Schule*, 158. Abgerufen von <https://www.forschungsmonitor-schule.de/print.php?id=122>

## Urheberrecht

Dieser Text steht unter der [CC BY-NC-ND 4.0 Lizenz](#). Der Name des Urhebers / der Urheberin soll bei einer Weiterverwendung wie folgt genannt werden: Patrizia Kühner (2021) für den [Forschungsmonitor Schule](#).