



Schulqualität sichern – Kompetenztests und Bildungsmonitoring



Im Überblick	4	Thema	6
Editorial: Bildung sichern	5	Qualität sichern im Südtiroler Bildungssystem	7
		Prove standardizzate nazionali e internazionali	9
		Viele Tests – ein Ziel	12
		Die PISA-Studie – Motor für die lokale und internationale Bildung	14
		Einzelsschule unter der Lupe	16
		Übersicht über die Kompetenztests	18



Service

30



Von den Taten vor den Daten	20	Frequently Asked Questions	31
Welche Daten braucht die Schule?	22	Buchtipps und Rezensionen	32
Externe Evaluation der Einzelschulen	24	Glossar	34
Qualitätsmanagement an der Berufsschule	25	Impressum	36
Kompetenztests – Chancen und Grenzen der Kompetenztests	26		
Das Messen des Nicht-Messbaren	28		

Im Überblick



Bildung sichern



Kompetenztests finden für alle Schülerinnen und Schüler statt, unabhängig von Schulstufe, Herkunft oder Schulleistung. Als solche garantieren sie auch Bildung für alle, weil auf Schwankungen auf breiter Basis reagiert werden kann. Sie dienen dem nationalen und internationalen Vergleich von Kompetenzen, die es zu erreichen gilt, damit Bildung für alle ermöglicht wird.

Die Evaluationsstelle ist verantwortlich für die Durchführung und die sinnvolle Aufbereitung der Daten in Form von Berichten oder Statistiken.

Testitems sind keine Lernaufgaben, um Unterrichtsinhalte, die gerade anstehen, zu überprüfen. Der einzelne Fachlehrer oder die einzelne Fachlehrerin wird Einblick in die Aufgaben nehmen und die Ergebnisse für seine und ihre Klasse interpretieren.

Schwindeln ist zwecklos – es wird statistisch errechnet, erkannt und korrigiert.

Auch Schulleitungen richten Strategien und Handlungen in der Schul- und Unterrichtsentwicklung an den Ergebnissen aus.

Für die Bildungspolitik stellen die regelmäßige Datenerhebung und die Trends eine wichtige Entscheidungsgrundlage dar. Die Kompetenzen in Mathematik und in den Sprachen finden sich in allen anderen Fächern und werden deshalb vorrangig getestet.

Das vorliegende Heft informiert über Tests im nationalen und internationalen Kontext, über Neuerungen in Südtirol und Italien, über Zusammenhänge zwischen Unterricht und Testergebnissen.

Die korrekte Information zu den Aufgaben, zur Durchführung, zur Auswertung und zur Datenverarbeitung ermöglicht es, einen professionellen Zugang zu den Kompetenztests zu finden. Damit alle davon profitieren und Bildung – im Sinne der Chancengleichheit – gesichert ist.

Die Leiterin Ursula Pulyer und die Mitarbeiterinnen und Mitarbeiter der Evaluationsstelle

Schulqualität sichern



Qualität sichern im Südtiroler Bildungssystem

Ein Konzept mit vielen Verbindungen

Grundlage für ein gutes, zeitgemäßes Bildungssystem ist das stete Streben nach Optimierung, nach Verbesserung, nach Entwicklung. Dies bedingt die Auseinandersetzung mit unterschiedlichen Qualitätsansprüchen (fachlichen und überfachlichen), mit angestrebten und tatsächlich erzielten Ergebnissen (Output und Outcome), mit der eigenen Leistungsfähigkeit (Performance). Dafür ist ein differenziertes Qualitätskonzept notwendig, in dem interne und externe Evaluation, Rahmenvorgaben und Unterstützungsangebote optimal ineinandergreifen. Was sind die Merkmale des Qualitätskonzeptes des deutschsprachigen Südtiroler Bildungssystems?

Merkmals 1

Qualität entsteht vor Ort und ist an der Schule zu sichern

Die Grundüberzeugung, Qualität systematisch und professionell an der Schule zu erheben, hat sich durchgesetzt; damit wird der internen Evaluation große Bedeutung beigemessen.

Bereits vor dem Inkrafttreten des Landesgesetzes zur Autonomie der Schulen im Jahre 2000 hat das damalige Pädagogische Institut die Schulen in ihrer Annäherung an die Evaluation begleitet und unterstützt. Daraus sind in den 1990er-Jahren zahlreiche Projekte entstanden, die in der Publikation „Spiegel aufstellen“ (1999) zusammengefasst sind.

Das Prinzip der **Rechenschaftslegung vor der eigenen Schulgemeinschaft** wurde im Jahre 2016 im **Dreijahresplan des Bildungsangebotes** verankert. Die Schulen sind seitdem verpflichtet, darin das eigene **Qualitätskonzept** zu beschreiben und zu **veröffentlichen**. Dieses ist bereits seit 2012 vorgesehen.

Auf Landesebene steht den Schulen seit vielen Jahren das Instrument **IQESonline** zur Verfügung, das sie neben **eigenen Instru-**

menten sehr effizient für unterschiedlichste Erhebungen – nicht nur zur Qualität – nutzen können. Die Evaluationsstelle stellt daneben auch **Leitlinien für die interne Evaluation** zur Verfügung.

Die Verknüpfung dieser Entwicklungsbestrebungen der Einzelschule mit jenen des Gesamtsystems erfolgt über die **Zielvereinbarung** zwischen der Schulführungskraft und dem Schulamtsleiter oder der Schulamtsleiterin.

Merkmals 2

Der systematische Außenblick fördert Qualität

Die Rechenschaftslegung in Form der internen und externen Evaluation ist als Verpflichtung für jede Schule und damit auch für das Gesamtsystem im Landesgesetz Nr. 12 aus dem Jahre 2000 grundgelegt. Durch dieses Gesetz erhalten die Schulen Autonomie in den Bereichen Organisation, Didaktik, Verwaltung, Forschung (Schulentwicklung) und Finanzen.

Die an der **Evaluationsstelle** angesiedelte **externe Evaluation** wurde deshalb ab dem Jahr 2004 aufgebaut und stetig weiterentwickelt. Im Rahmen der externen Evaluation

erstellt die Evaluationsstelle jährlich ein **Arbeitsprogramm**; den **Abschlussbericht** dazu bringt sie der Landesregierung und dem Landesschulrat zur Kenntnis. Sie führt neben **Schulbesuchen** – mit entsprechenden detaillierten Berichten an die Einzelschulen – **fokussierte Erhebungen, Umfragen** sowie **nationale und internationale Lernstandserhebungen** durch. Zu all diesen Erhebungen verfasst sie Landesberichte.

Die Auswertung, Aufbereitung und Vorstellung der Daten stellt für die Schulen und für die bildungspolitischen Verantwortsträger eine wertvolle Grundlage für die Qualitätssicherung und -entwicklung auf der Ebene der Einzelschule, aber auch auf der Ebene des Gesamtsystems dar. In den letzten Jahren entwickelte die Evaluationsstelle den partizipativen Ansatz und die Evaluation der Selbstevaluation (Metaevaluation) der autonomen Schulen, was sich in der Qualitätsentwicklung allgemein sehr bewährt hat.

Merkmals 3

Ein gemeinsames Qualitätsverständnis fördert Qualität

Die Grundlage für Unterricht und Didaktik bilden die **Rahmenrichtlinien** für den Kin-

dergarten, die Unterstufe, die Oberstufe und die Schulen der Berufsbildung. Ab dem Jahr 2008 wurden sie systematisch erarbeitet.

Der rote Faden, der sich durch alle Rahmenrichtlinien zieht, ist dabei die **Kompetenzorientierung**; die Weiterentwicklung der pädagogischen Arbeit und des Unterrichts gehen mit ihr einher.

Die Südtiroler Landesregierung hat im Jahr 2014 einen verbindlichen Qualitätsrahmen für die Schulen in Südtirol verabschiedet. Er ermöglicht eine klare Orientierung für die Gestaltung des Unterrichts und des Schullebens. Zu den Dimensionen wie **Inputqualitäten** (Kontext und Ressourcen), **Prozessqualitäten** (Lern- und Erfahrungsraum, Schulkultur und Schulklima, Schulführung, Professionalisierung und Schulentwicklung) und **Wirkungsqualitäten** (fachliche, überfachliche und erzieherische Kompetenzen) zeigt er Teilbereiche und Indikatoren auf, an denen die Schulen ihre Arbeit ausrichten. Die Schulbesuche im Rahmen der externen

Evaluation mit dem ausführlichen Bericht richten sich ebenfalls am Qualitätsrahmen aus.

Merkmal 4 **Unterstützung und Beratung** **fördern Qualität**

Gute, verlässliche und planbare **Rahmenbedingungen** erleichtern Qualitätsentwicklung und -sicherung erheblich. Dazu zählen neben den Ressourcen (Personal und Finanzen), neben den zeitgemäßen Infrastrukturen, den fachlich hochstehenden Fortbildungsangeboten, den adäquaten Unterrichtsmaterialien vor allem effiziente **Beratungs- und Unterstützungsangebote** zur Schul- und Unterrichtsentwicklung. Über die fünf **Pädagogischen Beratungszentren** im Bereich Innovation und Beratung (in Zukunft die Pädagogische Abteilung) können die Kindergärten und Schulen darauf zurückgreifen.

Das Qualitätskonzept, wie es hier beschrieben ist, hat sich in den letzten fünfzehn Jahren herauskristallisiert und es wird sich weiter entwickeln: Es bleibt ein Prozess. Das Qualitätskonzept zeigt, wie die Aufgaben aller Einrichtungen und Bereiche in unserem Bildungssystem miteinander verbunden sind, sie bauen aufeinander auf und entsprechend eng arbeiten sie zusammen. Wir alle suchen beständig nach Wegen dafür, wie wir eine gute Wirksamkeit unserer Kindergärten und Schulen erreichen können.

Peter Höllrigl

Schulamtsleiter und Ressortdirektor

Prove standardizzate nazionali e internazionali

La rivelazione degli apprendimenti di base

1. Introduzione

Gli ultimi decenni sono stati caratterizzati da profondi cambiamenti sociali ed economici, molti dei quali ancora in corso e dagli esiti molti difficili da prevedere e da interpretare. Tali processi hanno talvolta assunto la dimensione di veri rivolgimenti che hanno notevolmente modificato le prospettive di crescita e di sviluppo delle società avanzate, ponendo al centro dell'attenzione la qualità del capitale sociale di ciascun paese. È ormai evidente ai più che l'aggravamento della crisi economico-sociale degli ultimi anni richieda paradigmi nuovi per trovare possibili soluzioni, salvaguardando i traguardi fondamentali che le moderne economie avanzate hanno raggiunto nel secolo appena concluso (Wagner, Tony (2008)).

In questa prospettiva assume naturalmente un ruolo cruciale la qualità dell'istruzione e della formazione che ogni paese è in grado di garantire alla propria popolazione, sia a quella in età scolare sia a quella adulta. Valutare la qualità dell'istruzione è un tema molto complesso e richiede la considerazione di aspetti diversi e di natura differente. Tuttavia, è opinione diffusa, specie a livello internazionale, che qualsiasi valutazione

non possa prescindere da una solida misurazione degli esiti di apprendimento degli studenti, siano essi giovani studenti o adulti in formazione, di qualsiasi tipo essa sia. Tale istanza è divenuta ancora più rilevante nel momento in cui l'Italia, come molti altri paesi, ha adottato provvedimenti normativi volti alla riorganizzazione del sistema scolastico nazionale basandolo sull'autonomia delle singole istituzioni scolastiche. È, infatti, del tutto evidente che un sistema di scuole autonome richieda strumenti di valutazione e di misurazione per monitorare con attenzione che quantomeno le competenze di base siano a garantite a tutti e a ciascuno. Solo in questa prospettiva l'autonomia si può realmente tradurre in una concreta risorsa per il miglioramento e non trasformarsi, invece, in una potenziale fonte di disuguaglianze e iniquità all'interno del sistema educativo nazionale.

In questo contesto, già a partire dagli ultimi decenni del secolo scorso si sono affermate ricerche comparative internazionali¹ finalizzate alla misurazione dei livelli di competenza degli studenti della scuola primaria e secondaria in alcuni ambiti di disciplinari, ossia le cosiddette competenze di base fon-

damentali per l'esercizio dei diritti di cittadinanza attiva, principalmente riconducibili alla comprensione della lettura della lingua scritta, alla matematica e, talvolta, alle scienze naturali.

L'Italia ha partecipato quasi sempre alle ricerche internazionali sugli esiti di apprendimento, anche se è dovuto passare molto tempo affinché i loro risultati fossero posti al centro del dibattito nazionale sul sistema scolastico e divenissero un elemento di riferimento, anche se parziale, di azioni positive sul sistema scolastico nazionale. Inoltre, come è avvenuto in molti paesi, anche in seguito agli esiti delle ricerche precedentemente richiamate, l'Italia si è dotata di un proprio sistema nazionale per la rilevazione standardizzata degli apprendimenti di base², confluito recentemente in un più organico e completo Sistema Nazionale di Valutazione (SNV) del sistema scolastico italiano³.

2. I risultati italiani

Sin dalle loro prime edizioni, l'Italia partecipa a quasi tutte le rilevazioni internazionali sugli apprendimenti, con risultati non pienamente soddisfacenti, talvolta preoccupanti e forieri di ulteriori difficoltà in futuro. Gli

¹ Le più importanti e le più conosciute sono quelle promosse dalla IEA (International Association for the Evaluation of Educational Achievement) e dall'OCSE (Organizzazione per la Cooperazione e lo Sviluppo Economico).

² Le rilevazioni riguardano, al momento, la comprensione della lettura della lingua (italiana) scritta e la matematica (classi II e V della scuola primaria, classi I e III della scuola secondaria di primo grado e classe II della scuola secondaria di secondo grado), ma già a partire dall'anno scolastico 2013-14 è prevista un'estensione degli ambiti di competenza oggetto di rilevazione e dei livelli scolastici interessati.

³ Regolamento del sistema nazionale di valutazione in materia di istruzione e formazione approvato in via definitiva dal Consiglio dei Ministri nella seduta dell'08.03.2013.

esiti delle rilevazioni della IEA⁴ e dell'OCSE⁵ mettono in luce un quadro con molte ombre e poche luci, specie negli ambiti scientifici, ma non solo. Emerge con tutta evidenza che in un contesto comparativo internazionale il sistema scolastico italiano fatica a produrre buoni livelli di competenza, dimostrandosi quindi la necessità di disporre di misurazioni comparativamente solide per comprendere gli esiti effettivamente prodotti dalla scuola italiana.

I risultati del PISA (Programme for International Student Assessment), ma anche quelli del TIMSS (Trends in International Mathematics and Science Study) e del PIRLS (Progress in International Reading Literacy Study), hanno mostrato con tutta evidenza che il sistema italiano è caratterizzato da profonde differenze nelle sue articolazioni regionali, e che, in generale, vede al suo interno alte percentuali di allievi con livelli di competenza ritenuti a livello internazionale del tutto insufficienti e numeri piuttosto limitati di studenti eccellenti.

Questi risultati, tutt'altro che incoraggianti, permettono di affrontare su un piano probabilmente più adeguato il tema dell'equità del sistema scolastico italiano. Sovente molti esperti pongono comprensibilmente l'accento sul tema dell'equità, rilevandone la centralità e la rilevanza generale. Tuttavia, è opportuno osservare che in un sistema scolastico maturo l'equità del sistema scolastico non possa essere osservata solo

lungo la dimensione dell'inclusione nel sistema, ma anche in funzione dei livelli di apprendimento raggiunti dagli allievi. Ma a ben rifletterci, quest'ultimo aspetto non può essere verificato in assenza di un sistema di misurazione comparativamente solido e affidabile.

3. La rilevazione nazionale degli apprendimenti in Italia

A partire dagli ultimi anni del decennio appena trascorso l'Italia si è dotata di un servizio nazionale per rilevazione degli apprendimenti di base mediante prove standardizzate⁶. Esso mette a disposizione di ogni scuola, oltre che dell'intero sistema scolastico nazionale, anche nelle sue articolazioni regionali, dati comparativamente solidi e dettagliati. Inoltre, a partire dall'anno scolastico 2012-2013 l'INVALSI restituisce a ciascuna scuola gli esiti delle prove nazionali anche in una prospettiva di valore aggiunto, ossia al netto dell'effetto del contesto socio-economico-culturale in cui opera ogni istituzione scolastica (Ricci, Roberto (2008); Martini, Angela / Ricci, Roberto (2010), 80-107).

La realizzazione di un sistema nazionale per la rilevazione degli apprendimenti ha rappresentato nel panorama italiano un elemento di forte innovazione, certamente ancora incompleto e con alcuni aspetti che hanno bisogno di ulteriori aggiustamenti, in grado di fornire al Paese un'infrastruttura

immateriale per promuovere l'innalzamento dei livelli di competenza degli studenti italiani (Ricci, Roberto / Sestito, Paolo (2012)).

La costruzione del sistema nazionale di valutazione ha rappresentato e rappresenta una sfida sotto diversi punti di vista, non da ultimo quello metodologico-statistico.

Per la prima volta su scala nazionale sono state costruite prove secondo standard tecnico-scientifici adottati in ambito internazionale, basati su rigorose metodologie psicometriche e statistiche (Ricci, Roberto (2011), 21-23; Falzetti, Patrizia / Ricci, Roberto (2011a), 1-27).

Nonostante le rilevazioni realizzate dall'INVALSI siano oggetto di un acceso dibattito all'interno del sistema scolastico, raramente esso arriva a toccare realmente il contenuto delle prove e i criteri secondo i quali esse sono costruite. La costruzione di prove oggettive standardizzate⁷ è il frutto di un lungo e delicato processo, sovente non completamente conosciuto anche da molti degli utilizzatori delle prove stesse e dai diversi soggetti che operano nel mondo della scuola e della formazione in generale. Pochi realmente sanno che la formulazione di una prova standardizzata rivolta potenzialmente a centinaia di migliaia di studenti è l'esito di un lavoro profondamente e realmente interdisciplinare che coinvolge esperti con formazione ed esperienze specifiche e molto differenti tra di loro. Non sempre è noto che la costruzione di una prova standardizzata

⁴ Trends in International Mathematics and Science Study (rivolto alla IV classe della scuola primaria e alla III classe della scuola secondaria di primo grado) e Progress in International Reading Literacy Study (rivolto alla IV classe della scuola primaria).

⁵ Programme for International Student Assessment (rivolto agli studenti quindicenni negli ambiti: comprensione della lettura, matematica e scienze naturali).

⁶ L'attuale servizio nazionale di valutazione dell'INVALSI (Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione) è stato istituito nel 2007, dopo una fase sperimentale iniziata già nei primi anni 2000.

⁷ In questa sede per prove oggettive standardizzate si intendono prove costruite in modo che la loro correzione possa avvenire secondo modalità riproducibili e non dipendenti dal soggetto che effettua la correzione stessa.

richiede grandi sforzi e tempi piuttosto lunghi, mai inferiori ai 15–18 mesi, e il rispetto di una procedura molto articolata che sarà di seguito esposta nei suoi elementi essenziali.

Per favorire l'allargamento del dibattito a tutti gli aspetti connessi alle prove e al loro utilizzo l'INVALSI ha iniziato a promuovere eventi, trasmessi anche via web, per ampliare l'accesso alla discussione, anche critica, su temi così importanti e cruciali per lo sviluppo del sistema scolastico nazionale⁸.

4. Le prospettive future

I sistemi nazionali di valutazione sono, per definizione, strutture molto complesse e di non facile gestione. Essi richiedono continui aggiustamenti e potenziamenti per fornire al sistema scolastico, in tutte le sue articolazioni, informazioni solide e utili e, soprattutto, in grado di rispondere a istanze che cambiano rapidamente lungo diverse direttrici.

L'approvazione del regolamento istitutivo del Sistema Nazionale di Valutazione, finalizzato alla valutazione complessiva delle scuole e non solo degli apprendimenti, rappresenta un notevole impulso per l'intero processo di valutazione delle scuole al quale l'INVALSI dovrà fornire adeguate e convincenti risposte. Per limitare l'attenzione agli aspetti che maggiormente interessano questo contributo, ossia la rilevazione degli apprendimenti, l'INVALSI sta realizzando, anche con la collaborazione del mondo accademico, l'ancoraggio diacronico delle prove nazionali e l'ancoraggio tra le prove nazionali e quelli internazionali (Falzetti, Patrizia / Ricci, Roberto (2011b), 43). Ciò renderà possibile nel giro di un paio d'anni, e già a partire dal 2014, di valutare gli esiti degli apprendimenti degli studenti non soltanto in una prospettiva sezionale, ogni anno separatamente, ma anche longitudinale, potendo quindi cogliere linee evolutive e di sviluppo. In questo modo sarà pienamente a disposizione del Paese

una vera infrastruttura sulla quale basare, anche se ovviamente non in via esclusiva, linee d'intervento, fondate su un solido riscontro empirico, volte all'innalzamento dei livelli di competenza dei giovani italiani, favorendo quindi la crescita economica e sociale collettiva. Già da tempo l'istituto sta inoltre sperimentando la costruzione di nuove prove per la misurazione entro breve degli apprendimenti nell'ultimo anno della scuola secondaria di secondo grado e di altri ambiti disciplinari (ad esempio l'inglese e le scienze naturali) rispetto a quelli rilevati finora, ossia la comprensione della lingua scritta e della matematica.

Roberto Ricci

Dirigente di ricerca INVALSI
Responsabile Area prove nazionali

⁸ Si veda <http://www.invalsi.it/invalsi/istituto.php?page=eventi> (03.10.2017).

Riferimenti bibliografici

- Falzetti, Patrizia / Ricci, Roberto (2011a): I modelli della famiglia di Rasch nelle ricerche sugli apprendimenti. *La Matematica nella Società e nella Cultura*, Serie I, Vol. IV, 2011, 1–27.
- Falzetti, Patrizia / Ricci, Roberto (2011b), Un primo esperimento sullo studio diacronico dei risultati della prova nazionale INVALSI al termine del primo ciclo di istruzione in Italia. *Induzioni*, 21–34, 43.
- Martini, Angela / Ricci, Roberto (2010): Un esperimento di misurazione del valore aggiunto delle scuole sulla base dei dati PISA 2006 del Veneto. *Rivista di Economia e Statistica del territorio*, 2010, 3, 80–107.
- Ricci, Roberto (2008): La misurazione del valore aggiunto nella scuola. Collana Workingpaper della Fondazione Giovanni Agnelli, FGA, 9.
- Ricci, Roberto / Sestito, Paolo (2012): Il senso delle prove. *La voce.info*, <http://www.lavoce.info/archives/2706/il-senso-delle-prove/> (03.10.2017)
- Ricci, Roberto (2011): Le prove standardizzate, *L'Indice della scuola*, 12, III, 21–23.
- Wagner, Tony (2008): *The global achievement gap. Why Even Our Best Schools Don't Teach the New Survival Skills Our Children Need and What We Can Do About It*. Cambridge: Basic Books.

Lernstandserhebungen, Kompetenztests, INVALSI und Co.

Viele Tests – ein Ziel

Eine Vielfalt von Begriffen kursiert in der Schule im Zusammenhang mit der Überprüfung von erworbenen Kompetenzen von Schülerinnen und Schülern einer bestimmten Jahrgangsstufe: Kompetenztests, Vergleichsarbeiten, VerA-Arbeiten, INVALSI-Tests, Lernstandserhebungen. Wie auch immer die Tests bezeichnet werden – im Wesentlichen wird mit allen Testungen dasselbe Ziel verfolgt.

In den deutschen Schulen Südtirols werden seit mehreren Jahren regelmäßig Lernstandserhebungen geschrieben, die der Qualitätssicherung dienen und den Lehrkräften zusätzliche diagnostische Informationen zum individuellen, klassenbezogenen und schulweiten Kompetenzstand liefern. Auf gesamtstaatlicher Ebene ist das INVALSI (Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione) für die Durchführung und Auswertung der Lernstandserhebungen zuständig. Bedingt durch die besondere sprachliche Situation in Südtirol und ermöglicht durch die Zuständigkeiten der Autonomen Provinz Bozen wurden bestimmte Lernstandserhebungen (Deutsch und Mathematik 3. Grundschule und 1. Mittelschule) in Anlehnung an die in Deutschland länderübergreifend organisierten Vergleichsarbeiten (VerA) durchgeführt, aber auch Arbeiten des INVALSI übersetzt (Mathematik 3. Mittelschule und 2. Oberstufe) oder durch den Bereich Innovation und Beratung (BIB) und mit wissenschaftlicher Begleitung ausgearbeitet (Deutsch 3. Mittelschule). Im Schuljahr 2016/2017 wurden folgende Lernstandserhebungen in den Fächern Mathematik und Deutsch durchgeführt: **siehe Tabelle 1.**

Außerdem wurden in den vergangenen Jahren auch einzelne Kompetenztests in der Zweitsprache Italienisch und in Englisch durchgeführt oder pilotiert. Die Evaluationsstelle ist seit dem Schuljahr 2016/2017 für die Abwicklung aller Lernstandserhebungen zuständig.

Änderungen durch „La buona scuola“

Aufgrund staatlicher Vorgaben (Gesetz Nr. 107 vom 13.07.2015 „La buona scuola“; Legislativdekret Nr. 62 vom 13.04.2017) werden künftig auf Staatsebene Erhebungen in verschiedenen Schul- bzw. Klassenstufen in der Unterrichtssprache Italienisch, in Mathematik und Englisch verpflichtend eingeführt, wobei die Durchführung ab der Mittelschule am Computer erfolgt. Aufgrund der Sprachsituation in Südtirol wurde mit dem INVALSI vereinbart, dass in den deutschen Schulen Südtirols ab dem Schuljahr 2017/2018 die Lernstandserhebungen wie folgt durchgeführt werden: **siehe Tabelle 2.**

Die Arbeiten, die vom INVALSI kommen, werden ab der Mittelschule auch in Südtirol am Computer durchgeführt, sodass der

Korrekturaufwand für die Lehrpersonen deutlich reduziert wird; auch die VerA-Tests in Klassenstufe 8 werden digital abgewickelt, während jene für die 3. Grundschule und die 1. Mittelschule vorerst nur in Papierform zur Verfügung stehen. Von besonderer Bedeutung ist die Tatsache, dass die Teilnahme an den Erhebungen in den Abschlussklassen der Mittelschule und der Oberstufe die Voraussetzung für die Zulassung zur jeweiligen Abschlussprüfung darstellt.

Lernstandserhebungen versus Klassenarbeiten

Lernstandserhebungen unterscheiden sich wesentlich von Klassen- bzw. Schularbeiten und verfolgen grundsätzlich andere Ziele: Lernstandserhebungen mit ihren Testitems erfüllen psychometrische und inhaltliche Gütekriterien, dienen dem Nachweis von abgrenzbaren Kompetenzen einer definierten Schwierigkeit, sind klar hinsichtlich erwarteter Lösung, werden nach standardisierten Vorgaben korrigiert und liefern eine Rückmeldung für die Schul- und Unterrichtsentwicklung. Klassen- und Schularbeiten hingegen sind an den Lernstand einer Lerngruppe angepasst, dienen dem lang-

Tab. 1 – Lernstandserhebungen Schuljahr 2016/2017

Klassenstufe	Fach	Aufgabenentwicklung
3. Grundschule	Mathematik	VerA 3
	Deutsch	
1. Mittelschule	Mathematik	VerA 6
	Deutsch	
3. Mittelschule	Mathematik	INVALSI (Übersetzung ins Deutsche)
	Deutsch	BIB (in Anlehnung an Italienisch-Arbeit des INVALSI)
2. Oberstufe	Mathematik	INVALSI (Übersetzung ins Deutsche)

fristigen Erwerb von Kompetenzen, können unterschiedliche Lösungswege zulassen, weisen Unterschiede in der Bewertung auf und geben ein Feedback zum Lernprozess.

Pilotierung und Lösungshäufigkeit

Bevor Aufgaben bei einer Lernstandserhebung zum Einsatz kommen, werden sie einer sogenannten Pilotierung bzw. Feldstudie unterzogen, wo sie in Hinblick auf mehrere Kriterien überprüft werden. Das Kriterium der Validität bedeutet dabei, ob eine Aufgabe die zu überprüfende Kompetenz auch erfasst. Die Objektivität zielt auf die eindeutige Auswertbarkeit einer Aufgabe ab, unabhängig davon, wer die Korrektur vornimmt. Bei der Reliabilität geht es um die zuverlässige Erfassung eines Kompetenzbereiches. Zu guter Letzt stellt der Schwierigkeitsgrad einer Aufgabe ein wichtiges Kriterium dar. Dementsprechend kommen bei der Pilotierung verschiedene Aufgaben mit einem großen Schwierigkeitsspektrum zum Einsatz, mit dem Ziel, alle Kompetenzstufen zu berücksichtigen. Je weniger Schülerinnen und Schüler bei der Pilotierung eine Aufgabe lösen, desto schwieriger ist sie. Neben den sehr schwierigen Aufgaben, wo eine Lösungshäufigkeit von zehn Prozent erzielt

wird, gibt es auch sehr leichte Aufgabenstellungen mit einer Lösungshäufigkeit von 90 Prozent oder mehr. Die durchschnittliche Lösungshäufigkeit für ein Testheft liegt im Bereich zwischen 50 und 60 Prozent.

Aufgrund der überschaubaren Schülerpopulation Südtirols ist es möglich, dass die Pilotierungen ausschließlich durch externe Testleiter und Testleiterinnen durchgeführt werden. Die Ergebnisse einer Pilotierungsstudie werden den Schulen nicht rückgemeldet, da sie ausschließlich für die Testzusammenstellung vorgesehen sind.

„Vom Wiegen wird die Sau nicht fett“

Eine Schülerin oder ein Schüler wird vom Testen allein nicht schlau, Messungen und Tests allein bringen keine Verbesserung. Die Rückmeldungen zum Lern- und Leistungsstand der Schülerinnen und Schüler sagen isoliert nichts aus, was sich zur Weiterentwicklung des Unterrichts nutzen lässt. Wesentlich ist der schulinterne Umgang mit den Ergebnissen von Lernstandserhebungen. Erst wenn die Betroffenen und Beteiligten die Ergebnisse in ihre eigenen Sichtweisen und Erfahrungen von der Leistungsentwicklung, dem Unterricht und der

Schule integrieren, können Wirklichkeiten entstehen, aus denen sich sinnvoll Konsequenzen zur Weiterentwicklung des Unterrichts ableiten lassen. Erforderlich ist ein unaufgeregter und kompetenter Umgang mit den Rückmeldungen. Den Lehrkräften kommt dabei die Aufgabe zu, die Ergebnisse zu sichten und zu bewerten. Dies fällt vielen Lehrpersonen schwer, wird als zu zeitintensiv und nicht sinnvoll eingestuft. Nicht selten entfällt eine intensive Auseinandersetzung mit den Ergebnissen völlig. Mögliche Erkenntnisse zur Kompetenzentwicklung werden nicht genutzt. Für die schulinterne Auseinandersetzung mit den Ergebnissen braucht es die unvoreingenommene Bereitschaft von engagierten Lehrpersonen, die Daten als zusätzliche diagnostische Informationen anzunehmen und nach handlungsrelevanten Feldern zu suchen. Neben der Datenlesekompetenz müsste auch eine generelle Forschungskompetenz vermittelt werden, die Lehrkräfte befähigt, mit empirischen Befunden umzugehen und diese nicht als unhinterfragbare Wahrheiten anzunehmen.

Rosa Maria Niedermair
und Klaus Niederstätter

Evaluationsstelle für die deutsche Schule

Tab. 2 – Lernstandserhebungen ab Schuljahr 2017/2018

Klassenstufe	Fach	Aufgabenentwicklung
2. Grundschule	Mathematik	INVALSI (Übersetzung ins Deutsche)
3. Grundschule	Deutsch	VerA 3
5. Grundschule	Mathematik	INVALSI (Übersetzung ins Deutsche)
1. Mittelschule	Deutsch	VerA 6
3. Mittelschule	Deutsch	VerA 8
	Mathematik	INVALSI (Übersetzung ins Deutsche)
	Englisch	INVALSI
2. Oberstufe	Mathematik	INVALSI (Übersetzung ins Deutsche)
5. Oberstufe	Deutsch*	BIB (in Anlehnung an Italienisch-Arbeit des INVALSI)
	Mathematik*	INVALSI (Übersetzung ins Deutsche)
	Englisch*	INVALSI
4. Grundschule	Italienisch**	BIB
2. Oberstufe	Italienisch**	BIB

* ab Schuljahr 2018/2019 ** voraussichtlich ab Schuljahr 2018/2019

Die PISA-Studie

Motor für die lokale und internationale Bildung

Die PISA-Studie beeinflusst die Bildungspolitik der teilnehmenden Länder in unterschiedlichem Maße. Auch in Südtirol wurden Akzente gesetzt, die den Ergebnissen der Studie Rechnung tragen.

Mit der Durchführung der PISA-Studie im Jahr 2018 wird zum siebten Mal die Testung der fünfzehnjährigen Schüler und Schülerinnen unterschiedlicher Länder weltweit durchgeführt. Bei der ersten Auflage der Studie im Jahr 2000 haben in Summe 32 Länder teilgenommen. Inzwischen (2015) ist die Teilnehmeranzahl bereits auf 72 Länder angestiegen.¹ Die Trägerorganisation dieser groß angelegten Testung ist die OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung). Mit Sitz in Paris und unter der Schirmherrschaft des jeweiligen Generalsekretärs der OECD und dem Direktor des Direktorats für Bildung Andreas Schleicher wird diese internationale Studie zentral geplant und koordiniert.

Die Schwerpunkte

Verschiedene Grundbildungen stehen im Zentrum des Interesses von PISA. Die drei Hauptbereiche dabei sind das Lesen, die mathematische und naturwissenschaftliche Grundbildung. Wenn alle drei Jahre eine neue Auflage der PISA-Studie durchgeführt wird, gilt einem der drei Bereiche ein besonderes Interesse. In jenem Bereich werden vorab die Rahmenrichtlinien dafür nachjustiert und den Lernenden die meisten Aufgabenstellungen aus dem entsprechenden Bereich vorgelegt. Neben den drei obligatorischen Grundbildungen können sich die Länder entscheiden, ob sie noch eine oder beide fakultativen Grundbildungen – Finan-

cial literacy und Problem solving – dazunehmen möchten. Die Grafik auf der gegenüberliegenden Seite zeigt auf, welcher Bereich jeweils im Fokus stand bzw. steht.

Südtirol beteiligt sich seit dem Jahr 2003 mit einer repräsentativen Stichprobe an der PISA-Studie und ist demgemäß auch in den nationalen und internationalen Ergebnisberichten und Studien jeweils mit einer eigenen Auswertung der Ergebnisse präsent. Für die Autonome Provinz Bozen wurde seit dem ersten Jahr der Teilnahme ein eigener Ergebnisbericht entwickelt, in welchem die Gesamtergebnisse noch spezifischer nach Unterrichtssprachen der Schulen und Schultypen unterschieden wurden. In diesen Berichten wurde neben einer allgemeinen Darstellung auch die Auswertung zu den einzelnen Bereichen und zusätzlichen Faktoren vorgenommen. In den lokalen Berichten von 2012 und 2015 hat man damit begonnen, Trends herauszuarbeiten. Das war möglich, weil manche Aufgabenstellungen gleich waren und die Gesamtskalen für die Einzelbereiche gleichgeblieben sind. Im Unterschied zu den Vergleichsarbeiten bzw. Kompetenztests, die inzwischen in vielen Ländern zum Einsatz kommen, handelt es sich bei der PISA-Studie nicht um eine Vollerhebung mit der Fokussierung auf eine bestimmte Klasse und Schule. Es werden vielmehr repräsentativ in den teilnehmenden Ländern Stichproben gezogen, die über die statistische Hochrechnung

einen Einblick in die Bildungsentwicklung einer Provinz, einer Region oder einer gesamten Nation geben können. Hierbei steht also nicht die Mikro- und Mesoebene im Vordergrund, sondern die Makroebene eines Bildungssystems. Besondere Bedeutung erhält die PISA-Studie dann vor allem, weil sie in großen Teilen auf den Vergleich der verschiedenen teilnehmenden Länder ausgelegt ist – einem Vergleich, dem sich immer weniger Staaten zur Steuerung ihrer Bildungspolitik entziehen können.²

Ergebnisse beeinflussen Bildungspolitik

Diesbezüglich ist inzwischen eine umfangreiche Forschungsliteratur entstanden. Die tragende Frage der Forschenden in diesem Kontext ist: Wie stark ist der Einfluss der Ergebnisse auf die Bildungspolitik in den teilnehmenden Ländern? Dass ein „Effekt“ vorhanden ist, wird in der Fachliteratur nicht mehr bestritten, wenn auch die Richtung und die Stärke des Effekts unterschiedlich eingeschätzt werden. In einschlägigen Fachartikeln zum Thema wird versucht, die Felder des Einflusses der PISA-Studie abzu stecken. Drei Gruppen von Forschern und Forscherinnen können dabei ausgemacht werden: jene, die davon ausgehen, dass die Bildungspolitik durch PISA internationalisiert wurde, sodass einzelne Staaten kein Monopol mehr darüber haben; jene, welche ihr Augenmerk auf die Rolle verschiedener Erkenntnisse im bildungspolitischen Diskurs

¹ In einem übersichtlichen Artikel setzte sich z. B. Kerstin Martens (2007) mit den Wurzeln der PISA-Studie auseinander. Vgl. Die Pisa-Story. Wie es kam, dass die Bildungspolitik international wurde. Ein Lehrstück, in: Die Zeit 49/2007.

² Josef Schmid u. a. (2016) verweisen mehrmals in ihrem Werk: Internationalisierte Welten der Bildung. Bildung und Bildungspolitik im globalen Vergleich. Baden-Baden: Nomos Verlagsgesellschaft darauf, dass die PISA-Studie gemeinsam mit dem Bologna-Prozess als die Maßnahme schlechthin angesehen werden könne, die zu einer Internationalisierung der Bildung geführt habe. [Vgl. dazu S. 14, 79, 135 u. a.].

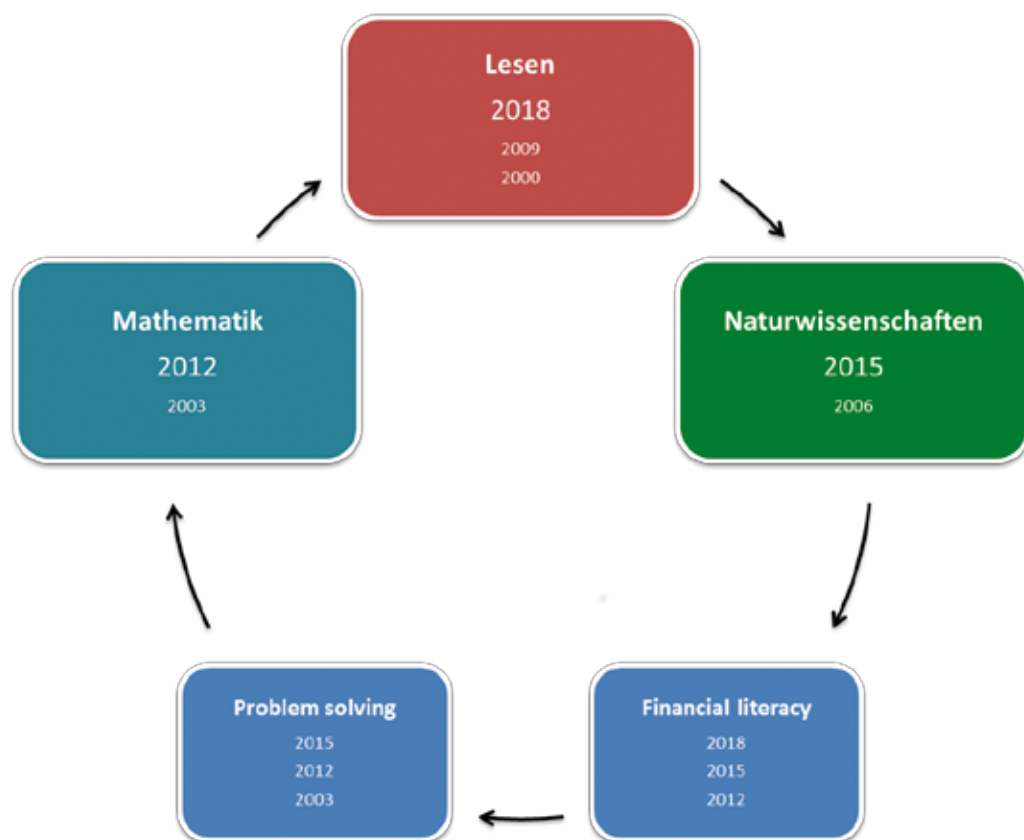


Abb. 1: Angeführt werden bei den Hauptbereichen jene Auflagen der PISA-Studie, in welchen diese Grundbildung im Fokus stand. Bei den fakultativen Bereichen werden die Jahre der Durchführung angeführt.

richten, und jene, welche explizit versuchen, die Rolle der OECD in der Globalisierung der Bildung zu analysieren.³ Welchen Effekt hatte die Studie auf die Bildungspolitik in Südtirol? Abgesehen von der Einführung der Rahmenrichtlinien für die deutschsprachigen Kindergärten und für alle Schulstufen im Zeitraum von 2008 bis 2012, die kompetenzorientiert⁴ gestaltet wurden, hat man in Südtirol, speziell für die deutsche Schule, eine Reihe von Maßnahmen von den Ergebnissen der PISA-Studie abgeleitet. Bei der Lesekompetenz hat man zum Beispiel das Augenmerk auf Schüler und Schülerinnen gerichtet, die mit ihren Ergebnissen die niedersten Kompetenzstufen erreicht haben. Zwei Aspekte wurden hierbei besonders

berücksichtigt: die Lesefertigkeit und die Lesefähigkeit. Diesbezüglich wurden verschiedene Kursfolgen und schulinterne Fortbildungskurse ins Leben gerufen und die Arbeit mit dem Werk „Lesen. Das Training“ gefördert. In diesem Zuge wurde auch das Programm „LUNA“ zur Früherkennung von Leseschwächen eingeführt und mit vielen Fortbildungen unterstützt. Im Bereich der Mathematik wurde ein Lehrmittel (Das Südtiroler Zahlenbuch, Stufe 1 bis 8) für den Unterricht zur Verfügung gestellt, welches den Bruch zwischen den verschiedenen Schulstufen verschwinden lassen sollte. Für die Mathematik und die Naturwissenschaften wurden spezielle Initiativen im MINT-Kontext gestartet, womit der großen Differenz in den

Ergebnissen zwischen den Mädchen und den Jungen entgegengewirkt werden sollte.⁵ In Summe sind bei den deutschen Schulen über die Lehrmittel, Unterstützungsprogramme und vor allem über ein reichhaltiges Fortbildungsprogramm Akzente gesetzt worden, die den verschiedenen Ergebnissen der PISA-Studie Rechnung tragen sollten. Was sich international in einigen Ländern zeigt, schlägt sich auch in der deutschen Schulwelt Südtirols nieder und eröffnet die Möglichkeit, den Sinn und die Bedeutung dieser internationalen Studie für unsere lokalen Gegebenheiten einzuordnen.

Ivan Stuppner

Evaluationsstelle für die deutsche Schule

³ Vgl. dazu Pons, Xavier (2017): Fifteen years of research on PISA effects on education governance. Eur J Educ. 52/2017, S. 133. Xavier Pons zeigt ebenso auf, dass die Studien zu den Einflüssen von PISA auf die Bildungspolitik in einigen Bereichen im Laufe der letzten Jahre zugenommen haben [S. 136–138]. In einem anderen Fachartikel versucht Sotiria Grek (2009) [Governing by Numbers: The PISA 'Effect' in Europe. Journal of Education Policy 24/2009, S. 14] den Einfluss der PISA-Studie anhand von einigen Ländern zu exemplifizieren, verweist schlussendlich auf den Begriff der Externalisierung nach Niklas Luhmann, welcher seiner Meinung nach die Reaktion der verschiedenen Länder in ihrer jeweiligen Bildungspolitik auf die PISA-Studie am besten zu charakterisieren vermag.

⁴ Für Südtirol hat Rudolf Meraner (2010) die Entwicklung der Kompetenzorientierung im Artikel: Kompetenzorientierung im Südtiroler Schulwesen, in: Erziehung und Unterricht 160/2010, H. 3-4, S. 374–383 thematisiert. In einem allgemeineren Kontext findet man bei Karl-Heinz Dammer in dessen Darstellung des Einflusses der PISA-Studie den Kompetenzbegriff als: „[...] kategorielles Bindeglied zwischen empirischer Bildungsforschung und politischer Steuerung“. Vgl. Dammer, Karl-Heinz (2015): Vermessene Bildungsforschung. Wissenschaftsgeschichtliche Hintergründe zu einem neoliberalen Herrschaftsinstrument. Baltmannsweiler: Schneider Verlag Hohengehren, S. 113–125.

⁵ Ein spezieller Dank gilt dem Bereichsleiter Rudolf Meraner für die Informationen zu den Maßnahmen für die deutschen Schulen hinsichtlich der PISA-Studie.

Einzelschule unter der Lupe

Schulqualität aus der Außenperspektive

Die Entwicklung der Einzelschule ist von standortbezogenen Daten abhängig. Lernstandserhebungen liefern schulspezifische Daten auf Basis einheitlicher und standardisierter Tests und sichern die Vergleichbarkeit der Ergebnisse.

Es gehört zu den Prämissen der Schulqualitätsicherung und -entwicklung, dass Schulleitungen eine Schlüsselrolle innehaben im Hinblick auf die Erfüllung gesellschaftlicher und bildungspolitischer Erwartungen an die Schule.¹ Um die gesellschaftlichen und innerschulischen Erwartungen in der schulischen Arbeit reflektieren zu können, ist eine Bewertung des Erreichten am Standort unabdingbar. Hierbei liefern Außenperspektiven wichtige Anstöße, die mit den Zielen der Schule in Zusammenhang stehen. Häufig wird durch den Blick von außen das Erreichte sichtbar und auch für den inneren Diskurs nutzbar. So relativierte sich in Österreich nach den ersten landesweiten Überprüfungen der Bildungsstandards in vielen Schulen eine zu pessimistische Eigenwahrnehmung. Ein Drittel der Schulleiterinnen und Schulleiter in Österreich gab nach Erhalt der ersten Rückmeldungen für die achte Schulstufe an, dass die Ergebnisse besser als erwartet waren. Für mehr als die Hälfte der Schulen bestätigten sich die eigenen Erwartungen. Lediglich für zehn Prozent fielen die Ergebnisse schlechter als erwartet aus. Eine fehlende Außenperspektive ging tendenziell damit einher, dass Erreichtes verkannt und eigene Leistungen zu schlecht eingeordnet wurden.²

Allerdings werden Wirkungen von Kompetenztests und Schulleistungsvergleichen häufig fälschlicherweise auf eine Perspektive der Rechenschaftslegung nach außen verkürzt. Hierbei trägt die große Aufmerksamkeit für internationale Schulleistungstudien, in Südtirol insbesondere PISA, dazu bei, dass der Nutzen von Kompetenztest in

der Sphäre von Schulaufsicht und Politik verordnet wird und der Wert einer Außenperspektive für die Einzelschule und ihre Akteurinnen und Akteure unterschätzt wird. Dabei sollte die Reflexion von Kompetenztests und Lernstandserhebungen von der Schule, ihren Zielen und Leitbildern her gedacht werden.

Ziele von Schule ernst nehmen – Grundkompetenzen absichern

In den zwei Jahrzehnten, nach den Schocks in vielen Ländern infolge der TIMSS-1995- und PISA-2000 Erhebungen, stand die Steuerung des Schulwesens unter dem Eindruck eines viel beschworenen Paradigmenwechsels: Der Schwerpunkt der Steuerung sollte sich von der Input- zur Output-Steuerung verlagern.³ Förderlich wäre, wenn verstärkt wahrgenommen würde, dass der scheinbar technokratischen Perspektive der Output-Steuerung eine zentrale Frage schulischen Handelns zugrunde liegen sollte: Welche Ziele soll Schule erreichen? Was erwarten Schülerinnen und Schüler, Eltern, Lehrerinnen und Lehrer sowie Gesellschaft von ihren Schulen? Die Ziele von Schule sind in vielfältigen Quellen für Akteurinnen und Akteure an den Schulen festgelegt, in Schulgesetzen, Curricula oder Schulleitbildern. In vielen Ländern wurden in den vergangenen Jahren Bildungsstandards festgelegt, die einen gesellschaftlichen Konsens über die zu vermittelnden Grundkompetenzen beschreiben sollen. Die mediale Rezeption von Schulleistungstests verkürzt deren Ergebnisse allerdings stark auf den Wettbewerb zwischen Ländern und Regionen in Form von Rangfolgentabellen. Doch gerade die

kritische Auseinandersetzung mit den eigenen Zielen führt von Schulleistungstests zu Anregungen für die Qualitätsarbeit.

Schulqualität aus der richtigen Perspektive betrachten

Für die Bewertung des Erreichten im Schulwesen gibt es viele Perspektiven. Jede Akteurin und jeder Akteur sollte sowohl Sicht auf den eigenen Wirkungsbereich als auch auf die Situation der anderen Ebenen des Systems haben. Aus der Perspektive Südtirols zeigen sich die Schulen im nationalen und internationalen Vergleich als erfolgreich in der Vermittlung grundlegender Kompetenzen. In den Naturwissenschaften und in Mathematik sind die in der PISA-2015-Testung gezeigten Kompetenzen jeweils circa ein halbes Lernjahr weiter fortgeschritten als im Schnitt der OECD-Länder.⁴ Dies zeigt sich auch im relativ geringen Anteil der Schülerinnen und Schüler, deren Kompetenzstand so gering ist, dass sie gefährdet sind, alltägliche Probleme und Situationen, auch im beruflichen Kontext, nicht bewältigen zu können: in Hinblick auf Mathematik 12,3 Prozent der Jugendlichen. Auch im Lesen zeigt die internationale Perspektive für Südtirol gute Leistungen, jedoch weniger deutlich. Hier finden sich auch mehr Jugendliche, deren geringe Lesekompetenz eine gesellschaftliche Teilhabe gefährden könnte.

Die Doppelperspektive des normativen Vergleichs, das heißt mit einer externen Gruppe (Länder, Regionen, Schulen) und des Bezugs auf ein kriteriales Ziel, hier dem Kompetenzstand der Jugendlichen relativ zu einer inhaltlichen Beschreibung desselben,

¹ Wissinger, Jochen (2016): Schulleitung im Fokus des Schulqualitätsdiskurses. In: Steffens, Ulrich / Bargel, Tino (Hgg.): Schulqualität – Bilanz und Perspektiven. Münster: Waxmann, S. 257–276.

² Rieß, Cornelia / Zuber, Julia (2014): Rezeption und Nutzung von Ergebnissen der Bildungsstandardüberprüfung in Mathematik auf der 8. Schulstufe unter Berücksichtigung der Rückmeldemoderation. Salzburg: BIFIE.

³ Altrichter, Herbert / Maag Merki, Katharina (Hgg.): Handbuch Neue Steuerung im Schulsystem. Educational Governance, Band 7. 2., überarbeitete und aktualisierte Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.

⁴ Evaluationsstelle für die deutsche Schule der Autonomen Provinz Bozen (2017). PISA 2015 – Ergebnisse Südtirol. http://www.provinz.bz.it/evaluationsstelle-deutschsprachiges-bildungssystem/download/PISA_2015_Ergebnisse_Suedtirol.pdf (09.10.2017)

eröffnet Raum für einen Diskurs. So wie sich für Südtirol zum einen die Reflexion darüber anbietet, was gut gelingt, fördert die Betrachtung des Erreichten mit Blick auf die angestrebten Zielkriterien die Weiterentwicklung des Systems.

Für die Entwicklung bieten internationale und nationale Daten jedoch nur einen Referenzwert. Kompetenzmessungen zeigen, dass die Unterschiede innerhalb eines Schulsystems, zwischen Schulen, weit größer sind als zwischen Ländern. Trennen Österreich und Südtirol in PISA im Lesen 20 Punkte, so liegt zum Beispiel innerhalb Österreichs zwischen starken und schwachen Schulen eine mehr als fünffache Spanne. Im Vergleich von Schulen mit ähnlichen Rahmenbedingungen und sozialer Zusammensetzung der Schülerinnen und Schüler liegen in Österreich zwischen Schulen mit sehr guter und geringer Zielerreichung 80 bis 100 Punkte.⁵ Obwohl sich die Schulen in ihrem Kontext von außen betrachtet nicht wesentlich unterscheiden, unterscheidet sich das mittlere Kompetenzniveau von Schulen gleicher Rahmenbedingungen im Ausmaß mehrere Lernjahre.

In Anbetracht der breiten Streuung der beobachteten mittleren Kompetenzen an den Schulen wird klar: Nur standortbezogene Daten können die Entwicklung an der Einzelschule fördern. Lernstandserhebungen wie INVALSI in Südtirol, die Kompetenztests VerA in Deutschland und Südtirol oder die Überprüfung der Bildungsstandards in Österreich bieten schulspezifische Daten auf Basis jeweiliger einheitlicher Tests, die unter standardisierten Bedingungen durchgeführt werden. Dies sichert die Vergleichbar-

keit der Testergebnisse. Genauso wichtig ist jedoch auch die individuelle Interpretation und Aufarbeitung der Ergebnisse am Schulstandort. Diese geschieht unter verschiedenen Blickwinkeln, wobei neben der zentralen Perspektive der Erreichung der Grundkompetenzen und der Ziele des Unterrichts die folgenden denkbar sind:

- **Verständnis der Rahmenbedingungen der Schule:** Idealerweise ergänzen Fragen zu den Eigenschaften der Schülerinnen und Schüler die Kompetenztests, zum Beispiel nach der gesprochenen Sprache, dem Geschlecht oder der Bildung der Eltern. Als aussagekräftig zeigt sich die Frage nach dem Buchbesitz in der Familie. Bei Rückmeldung der Daten können Schulen ihre Wahrnehmung der Rahmenbedingungen extern validieren und den Lernstand einordnen. In Österreich wird durch die Einbeziehung der Rahmenbedingungen ein Vergleich von Schulen in ähnlicher Lage möglich, der vielfach als wichtigstes Element des Feedbacks gesehen wird.
- **Verständnis der Leistungsstreuung:** Lernstandserhebungen ermöglichen es, die breite Streuung der Leistungen in einer Klasse bzw. Schule zu erfassen. Für Österreich zeigt sich, dass sich in allen Klassen Schülerinnen und Schüler finden, deren Leistungen so weit über oder unter dem mittleren Niveau der Klasse liegen, dass ein Unterricht, der auf dieses Niveau abzielt, sie kaum erreichen kann. In Österreich liegt circa ein Drittel aller Schülerinnen und Schüler mehr als ein Lernjahr vor dem mittleren Lernstand ihrer Klasse und gleichzeitig ein weiteres Drittel ein

Lernjahr hinter dem Stand zurück.

- **Betrachtung des Schulklimas und lernförderlicher Umstände:** Die Schulqualität schlägt sich auch in wesentlichen Rahmenbedingungen des Unterrichts und Lernens nieder. Entsprechende im Rahmen einer Kompetenzmessung erhobene Daten können das Bild der Schulqualität ergänzen.
- **Betrachtung von Gruppen:** Die Rückmeldungen zu externen Lernstandserhebungen erleichtern die Entwicklung verschiedener Schülergruppen, zum Beispiel von Mädchen und Jungen, im Überblick zu erfassen und zu thematisieren.
- **Verständnis von fachdidaktischen Mustern:** Lernstandserhebungen decken die angestrebten Grundkompetenzen in der ganzen Breite des Faches ab. Schulen können in ihren Leistungsprofilen individuelle Stärken und Schwächen des Unterrichts erkennen, zum Beispiel für mathematische Inhalte oder Prozesse. Werden Kompetenzen für mehrere Fächer zeitgleich gemessen, zeigt sich, ob Stärken fächerübergreifend oder begrenzt auf einzelne Bereiche vorliegen.
- **Zeitliche Entwicklung:** Sichtbar wird die Entwicklung am Standort, wenn vergleichbare Lernstandserhebungen im Abstand von mehreren Jahren wiederholt werden.

Michael Bruneforth

Berater im Bildungsbereich und stellvertretender Leiter des Departments Bildungsstandards & Internationale Assessments am Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens

⁵ Breit, Simone / Bruneforth, Michael / Schreiner, Claudia [2017]: Standardüberprüfung 2016 – Deutsch, 8. Schulstufe. Bundesergebnisbericht. Salzburg: BIFIE.

Übersicht über die

	DEUTSCH	ENGLISCH
In welchen Klassenstufen werden sie durchgeführt?	3. Klasse Grundschule 1. und 3. Klasse Mittelschule 5. Klasse Oberschule (ab 2018/2019)	3. Klasse Mittelschule 5. Klasse Oberschule (ab 2018/2019)
Woher kommen die Testaufgaben?	für Grund- und Mittelschule: Übernahme der Kompetenztests VerA 3, VerA 6 und VerA 8 aus Deutschland, entwickelt am IQB (Institut zur Qualitätsentwicklung im Bildungswesen) der Humboldt-Universität Berlin; für die Oberschule: Entwicklung am BIB (Bereich Innovation und Beratung)	Übernahme der Tests des INVALSI auf Staatsebene
Wie werden die Tests abgewickelt?	in der Grund- und Mittelschule: in Papierform, bis die Tests auch in Deutschland digital durchgeführt werden; für den Bereich Zuhören auch Einsatz von Audiodateien in der Oberschule: in der Anfangsphase voraussichtlich in Papierform	digital am Computer
Sind die Tests bewertungs- bzw. prüfungs-relevant?	keine Bewertung; Beteiligung am Test ist in der 3. Mittelschule und in der 5. Oberschule Voraussetzung für die Zulassung zur jeweiligen Abschlussprüfung	keine Bewertung Beteiligung am Test ist in der 3. Mittelschule und in der 5. Oberschule Voraussetzung für die Zulassung zur jeweiligen Abschlussprüfung
Was wird überprüft?	in der Grund- und Mittelschule: jeweils zwei Testdomänen, und zwar Lesen (jedes Jahr) und alternierend einer der folgenden drei Bereiche: Zuhören, Orthografie, Sprache und Sprachgebrauch in der Oberschule: in Anlehnung an die Aufgabenstellung des INVALSI für Italienisch	Kompetenz nach dem Europäischen Referenzrahmen (Mittelschule: Niveau A2; Oberschule: Niveau B2) in den Testdomänen Lesen (Reading), Zuhören (Listening) sowie Sprache und Sprachgebrauch (Language in Use)
Welche Aufgabentypen kommen vor?	in der Grund- und Mittelschule: Ankreuzaufgaben, Richtig-Falsch-Aufgaben, Aufgaben mit eigener Antwort / Kurzantwort, Zuordnungsaufgaben, Aufgaben mit Lückentext, vorgegebene Angaben in die richtige Reihenfolge bringen, Korrekturen vornehmen in der Oberschule: in Anlehnung an die Aufgabenstellung des INVALSI für Italienisch	Ankreuzaufgaben (Multiple choice), Richtig-Falsch-Aufgaben (right-wrong-justification), Aufgaben mit eigener Antwort / Kurzantwort (short answer questions), Zuordnungsaufgaben (matching), Aufgaben mit Lückentext (gap filling), vorgegebene Angaben in die richtige Reihenfolge bringen (Sequencing), Hörverständnisaufgaben (Listening comprehension)

Kompetenztests

MATHEMATIK	ITALIENISCH
2. und 5. Klasse Grundschule 3. Klasse Mittelschule 2. Klasse Oberschule 5. Klasse Oberschule (ab 2018/2019)	4. Klasse Grundschule (ab 2018/2019) 2. Klasse Oberschule (ab 2018/2019)
für Grund-, Mittel- und Oberschule: zum Einsatz kommen die vom INVALSI entwickelten und in Südtirol übersetzten Tests	von der Universität Siena in Kooperation mit dem Bereich Innovation und Beratung
in der Grundschule: in Papierform in der Mittel- und Oberschule: digital am Computer	in Papierform zukünftig computerbasiert angestrebt
keine Bewertung Beteiligung am Test ist in der 3. Mittelschule und in der 5. Oberschule Voraussetzung für die Zulassung zur jeweiligen Abschlussprüfung	keine Bewertung
in der 2. Grundschule: drei mathematische Bereiche: Zahlen, Ebene und Raum, Daten und Vorhersagen ab der 5. Grundschule: vier mathematische Bereiche: Zahlen, Ebene und Raum, Daten und Vorhersagen, Relationen und Funktionen anzuwendende mathematische Prozesse: Erkennen, Problemlösung, Argumentation	Kompetenzen orientiert am Europäischen Referenzrahmen in den Testdomänen Lesen (lettura), Zuhören (ascolto) sowie Sprache und Sprachgebrauch (lingua in uso)
in der Grund-, Mittel- und Oberschule: Ankreuzaufgaben, Richtig-Falsch-Aufgaben, Aufgaben mit eigener Antwort / Kurzantwort, Zuordnungsaufgaben, Aufgaben mit Lückentext, vorgegebene Angaben in die richtige Reihenfolge bringen, Korrekturen vornehmen	Ankreuzaufgaben (risposta multipla), Richtig-Falsch-Aufgaben (vero-falso), Aufgaben mit eigener Antwort / Kurzantwort (risposta breve), Zuordnungsaufgaben (abbinamento), Aufgaben mit Lückentext (completamento testo), vorgegebene Angaben in die richtige Reihenfolge bringen (ordine di sequenza), Hörverständnisaufgaben (comprensione orale)

Anmerkungen zum evidenzbasierten Schulleitungshandeln

Von den Taten vor den Daten

1. Die Einführung

„Neuer“ Steuerungskonzepte

Seit geraumer Zeit gehen mit der bildungspolitisch motivierten Einführung des Begriffes der „Neuen“ Steuerung schulsystemimmanente Veränderungen einher, die durch eine stärkere Ergebnis- und Wettbewerbsorientierung zum Ausdruck kommen.¹ Dabei geht es im Kern um die sogenannten vier „E“, um Effektivität, Effizienz, Erfolg und Evidenz. Hinzu kommt ein neues Führungsverständnis an den Schulen. Schulleitungen erhalten mehr Entscheidungskompetenzen und sollen in diesem Zusammenhang Prozesse einer schulbezogenen Strategie- und Organisationsentwicklung möglichst eigenverantwortlich initiieren, vorantreiben bzw. weiterentwickeln.²

Diesem von staatlicher Seite an die Schulen herangetragenen Zielsystem liegt ein Steuerungsdenken zugrunde, in dessen Zentrum die „Idee der Rationalisierung der Systemsteuerung des Bildungswesens durch die Verbesserung der Informationslage über seinen Entwicklungsstand und seine Auswirkungen“³ steht. Es sind die Erfolgskriterien einer solchen Rationalisierung zu definieren und sodann die Wirkungen der derart veränderten Systemsteuerung festzustellen. Mit der Evidenz, also dem Aufzeigen von Ergebnissen, dem Sichtbarmachen der Systemleistung, des sogenannten Outputs, als einem zentralen Element neuer Steuerungsannahmen annonciert das Konzept der „Neuen“ Steuerung eine Datenbasis, die Entscheidungsträgern auf den unterschiedlichsten Steuerungsebenen des Schulsystems Orientierungs- und Ent-

scheidungshilfe sein soll.⁴ Eine solche hohen methodischen Standards genügende Basis bietet den jeweiligen Entscheidungsträgern zum einen Anhaltspunkte für kritische Selbstreflexion, zum anderen ermöglicht sie ihnen bei der Entwicklung und Implementierung von Strategien, Zielsetzungen und Maßnahmen, mithin beim Management in ihrem Verantwortungsbereich, ein neues, hochinformiertes Agieren.⁵

Mit der Schulstatistik, die auf der Mikroebene sozioökonomische Rahmenbedingungen von Schulen markiert, sodann mit Datenbeständen aus großräumigen Untersuchungen zu unterschiedlichen Leistungsaspekten in Lernbiografien stehen für dieses Agieren wichtige Hilfsmittel zur Verfügung. Hinzu kommt, dass die Bildungspolitik insbesondere über Verfahren der externen Evaluation dazu beiträgt, die Datenbasis zu verbessern.⁶

2. Schulleitung als Adressatin der „Neuen“ Steuerungskonzepte

Die Untersuchung schulischer Entwicklungsstände soll weniger eine interessierte Öffentlichkeit, auch nicht ein spezifisches Schulpublikum bedienen, sie findet nicht zum Zweck eines Rankings statt, sondern soll zuerst Leitungen der Bildungseinrichtungen zugute kommen. Es sollen auf Grundlage der bereitgestellten Informationen mögliche Gründe für eventuell unbefriedigende Ergebnisse analysiert und geeignete Maßnahmen durch die mit der Gestaltung der Schule befassten Manager abgeleitet werden. Dieser datenbasierte Steuerungsanspruch geht über ein

Gestaltungsmanagement, das allein vom tradierten beruflichen Erfahrungs- und Wissenshorizont von Schulleiterinnen und Schulleitern getragen ist, hinaus. Der datenbasierte Steuerungsanspruch geht mit der Annahme einher, dass das Handeln der Protagonisten im Bildungssystem effektiver und effizienter würde, wenn ihnen mehr und besseres „Steuerungswissen“ zur Verfügung stünde.⁷ Die Schulleiterinnen und Schulleiter werden von daher dazu angehalten, die Befunde zu rezipieren, sie angemessen zu interpretieren und abschließend eine Ableitung sach- bzw. problemorientierter, bedarfsgerechter und adressatenadäquater Entscheidungen zu treffen, die ihrerseits in entsprechende Maßnahmenprogramme Eingang finden. Die darauf folgende Implementierung jener Maßnahmenpakete soll dann, der neuen Steuerungslogik folgend, mittelfristig zu einer Verbesserung des schulischen Outputs (z. B. fachliche Schülerleistungen) beitragen.

3. Zum Führungshandeln ist es ein langer und kurvenreicher Weg ...

Seit einiger Zeit wird das Modell der evidenzbasierten Steuerung im Sinne von evidence based policy/practice nun doch, so selbstverständlich und selbstregelnd es scheinen mag, einer kritischen Reflexion unterzogen. Die Frage nämlich, inwiefern die in verschiedenen Datenquellen enthaltenen Informationen über Bildungseinrichtungen für Steuerungshandlungen genutzt werden, ist bisher noch nicht hinreichend theoretisch untersetzt oder empirisch untersucht worden.⁸

¹ Vgl. Böttcher, Wolfgang (2016): Zur Kritik der pädagogischen Ökonomie – zum Verhältnis von Monitoring und Unterstützung. In: Steffens, Ulrich / Bargel, Tino (Hgg.): Schulqualität – Bilanz und Perspektiven. Grundlagen der Qualität von Schule. Münster und New York: Waxmann, S. 221–234.

² Vgl. Brauckmann, Stefan (2016): Schulleitungshandeln zwischen erweiterten Rechten und Pflichten (SHaRP) unter Berücksichtigung einschlägiger schulrechtlicher Bestimmungen. In: Bundesministerium für Bildung und Forschung (BMBF) Referat für Bildungsforschung (Hrsg.): Steuerung im Bildungssystem. Implementation und Wirkung neuer Steuerungsinstrumente im Schulwesen, 43. Band. Rostock: W. Bertelsmann Verlag GmbH & Co. KG, S. 238–255; Avenarius, Hermann / Brauckmann, Stefan / Döbert, Hans / Isermann, Kathrin / Kimming, Thomas / Seeber, Susan (2006): Durch größere Eigenverantwortlichkeit zu besseren Schulen. Ergebnisbericht der wissenschaftlichen Begleitung des „Modellvorhabens eigenverantwortliche Schule (MeS)“ im Land Berlin 1. Auflage. Berlin/Frankfurt am Main.

³ Altrichter, Herbert / Heinrich, Martin (2006): Evaluation als Steuerungsinstrument im Rahmen eines „neuen Steuerungsmodells“ im Schulwesen. In: Böttcher, Wolfgang / Brohm, Michaela / Holtappels, Hans-Günter (Hgg.): Evaluation im Bildungswesen. Weinheim: Juventa, S. 51–64 (Hier: 55).

⁴ Vgl. Döbert, Hans / Weishaupt, Horst (Hgg.) (2015): Bildungsmonitoring, Bildungsmanagement und Bildungssteuerung in Kommunen. Ein Handbuch. Münster/New York: Waxmann Verlag.

⁵ Vgl. Dederich, Kathrin (2016): Entscheidungsfindung in Bildungspolitik und Bildungsverwaltung. In: Altrichter, Herbert / Maag Merki, Katharina (Hgg.):

Handbuch Neue Steuerung im Schulsystem. Educational Governance, Band 7. 2., überarbeitete und aktualisierte Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 53–73.

⁶ Vgl. Weishaupt, Horst (2007): Überlegungen zur Konzeptualisierung einer Schulberichterstattung in Hessen für die Handlungsebenen Schule, regionales Schulum und Land [Gutachten]. Wuppertal.

⁷ Vgl. Altrichter / Heinrich (2006): Evaluation als Steuerungsinstrument im Rahmen eines „neuen Steuerungsmodells“ im Schulwesen.

⁸ Vgl. Dederich (2016): Entscheidungsfindung in Bildungspolitik und Bildungsverwaltung.

Folglich ist auch nur wenig darüber bekannt, inwieweit die Schulleiterinnen und Schulleiter die datengestützten Informationen rezipieren und im Sinne einer praktisch relevanten Umsetzung für Steuerungshandlungen nutzen. Vielmehr stehen noch weitestgehend Erkenntnisse darüber aus, ob und auf welche Weise Daten überhaupt als Instrumente der Führung und Entwicklung von Schule fungieren, ob sie einen Mehrwert für deren Leiterinnen und Leiter aufweisen und ob das bereitgestellte Wissen tatsächlich auch in schulische Führungs- und Entwicklungsaktivitäten mündet. Der Blick richtet sich damit auch auf die Bildungsverwaltung als Datenerzeuger. Sie muss sich vergegenwärtigen, dass sich die Konkretisierung des neuen Steuerungsansatzes den Schulleiterinnen und Schulleitern mitunter nur bedingt erschließt. Es wäre deshalb zu prüfen, wie die Schulleiterinnen und Schulleiter die staatlicherseits intendierten bzw. realisierten konzeptionellen Vorstellungen evidenzbasierten Führungshandelns bewerten und für wie angemessen sie diese für die Durchsetzung ihres Führungsanspruchs halten. Somit wäre zu klären, auf welche Weise das Konzept des evidenzbasierten Steuerungsmodells auf das evidenzbasierte Führungshandeln heruntergebrochen werden kann, so dass mit den Daten gearbeitet werden kann.

Darüber hinaus wäre es wichtig zu erfahren, inwiefern zum Zwecke der Verarbeitung des datengestützten „Steuerungswissens“ in- zwischen Veränderungen in der schulischen Leitungsstruktur vorgenommen wurden und welche Formen der inner- und außerschu-

lischen Kooperation dabei möglicherweise auftreten können bzw. müssen.

Als problematisch kann zudem gelten, dass sich der programmatische Anspruch datengestützten Leitungshandelns als Anreizhilfe für erfolgsorientiertere Steuerungshandlungen auf alle Schulen gleichermaßen bezieht, unabhängig davon, welchen sozioökonomischen und soziokulturellen Kontextbedingungen diese unterliegen.⁹

4. Herausforderungen an eine schul(leitungs)bezogene Datenrückmeldung ...

Nicht alle schulbezogenen Daten, die gegenwärtig erfasst werden können, sind als Daten für eine Evaluation qualitätsbezogenen Führungshandelns anzusehen und in diesem Sinne brauchbar und nutzbar. Dies würde unweigerlich zu Datenfriedhöfen führen. Vielmehr ist zu prüfen, welche unter der Vielzahl möglicher Indikatoren ausgewählt werden sollen, die auch auf die qualitätsbezogenen Wirkungsdimensionen des Schulleitungshandelns bezogen werden können. Dies wirft die Frage nach einem bewährten Analyserahmen für die Beschreibung wirkungsbezogenen Schulleitungshandelns auf. Eine stärkere Berücksichtigung des einzelschulischen Kontexts könnte Schulleiterinnen und Schulleitern in jedem Fall dabei helfen, den identifizierten Beratungsbedarf besser nachzuvollziehen und ihre Entscheidungen über Prioritäten der Organisations-, Personal- und Unterrichtsentwicklung kontextsensitiv zu begründen und gegebenenfalls umzusetzen.¹⁰ Ziel dieses Teils einer stärker kontextbezogenen schul(leitungs-)

bezogenen Datenrückmeldung müsste es daher sein, relevante nationale und regionale Rahmenbedingungen der Schul- und Unterrichtsentwicklung auszuwerten¹¹, die Stärken zu benennen und zugleich auf Bereiche hinzuweisen, die im nationalen und regionalen Vergleich als prioritäre Gestaltungsfelder der Qualitätsentwicklung im Schulsystem angesehen werden können.¹² Für das Konzept einer schulleitungsbezogenen Datenrückmeldung stellt sich zukünftig die Frage, ob mit den derzeit verfügbaren Daten überhaupt valide und aussagekräftige Indikatoren für alle nachgefragten Qualitätsbereiche und Dimensionen schulleiterischen Führungshandelns vorliegen, also Indikatoren, die handlungsrelevante Dimensionen der Schulführung in zentralen Aspekten der Organisations-, Personal- und Unterrichtsentwicklung auch erfassen.

Stefan Brauckmann

Professor für Qualitätsentwicklung und Qualitätssicherung im Bildungsbereich an der Alpen-Adria-Universität Klagenfurt

⁹ Vgl. Schwarz, Alexandra / Brauckmann, Stefan (2015): Between Facts and Perceptions: The Area Close to School as a Context Factor in School Leadership. Wuppertal, Germany: University of Wuppertal.

¹⁰ Vgl. Brauckmann, Stefan / Böse, Susanne (2017): Picking up the pieces? Zur Rolle der Schulleitung beim Turnaround – Ansätze und empirische Erkenntnisse. In: Maniti, Veronika / Dobbelstein, Peter (Hgg.): Schulentwicklungsarbeit in herausfordernden Lagen. Münster: Waxmann, S. 85–103.

¹¹ Vgl. Kühne, Stefan (2017): Stadt, Land, Muss? Notwendigkeit und Möglichkeiten der Regionalisierung von Bildungsindikatoren. Berufsbildung, 163 (71), S. 18–20.

¹² Vgl. Weishaupt (2007): Überlegungen zur Konzeptualisierung einer Schulberichterstattung in Hessen für die Handlungsebenen Schule, regionales Schulumfeld und Land.

Welche Daten braucht die Schule?

Quantitativ oder qualitativ

Die empirische Sozialforschung bietet zwei Wege, zu abgesicherten Erkenntnissen über das Schulwesen zu gelangen: den qualitativen und den quantitativen Weg. Während quantitative, statistische Ergebnisse mittlerweile einen festen Platz in der Sozialberichterstattung haben, sind qualitative Verfahren erst in Ansätzen bekannt. Worin unterscheiden sich die beiden Vorgehensweisen und welchen Beitrag können qualitative Methoden bei der Evaluation des Schulwesens leisten?

Ausgangspunkt für eine Unterscheidung ist zunächst das Erkenntnisinteresse: Was soll untersucht werden? Im Rahmen von Evaluationen ist dies stets verknüpft mit einem Verwertungsinteresse, das heißt Wissen soll so produziert werden, dass es Öffentlichkeit, Politik und Professionelle bei Entscheidungen zur Steuerung und Planung zum Beispiel im Schulwesen unterstützen kann.¹ Dies sind Fragen wie: Wie hoch ist das Durchschnittsalter der Schülerinnen und Schüler, die das Abitur absolvieren? Besteht ein Unterschied je nach sozialer Herkunft in den Lesekompetenzen? Wie gehen pädagogische Fachkräfte mit hohen Leistungsunterschieden im Klassengeschehen um? Welche Bedingungen helfen Schülerinnen und Schüler, Prüfungsdruck zu bewältigen?

Quantitativ heißt, es können zum Teil enorme Datenmengen wie bei PISA übersichtlich präsentiert und Zusammenhänge zwischen mindestens zwei Merkmalen für ganze Bevölkerungsgruppen analysiert werden. Interessant sind diese Daten vor allem im Vergleich, zum Beispiel zwischen Ländern, zwischen Schulformen, zwischen Alterskohorten. Diese Art der Zusammenfassung und Vergleichbarkeit erfordert eine Standardisierung, weswegen eine sehr konkrete Vorstellung über den Untersuchungsgegenstand benötigt wird.² Während eine Frage nach dem Alter oder nach dem (angestrebten) Bildungsabschluss keine großen Schwierigkeiten bereitet, ist dies bei einer Frage nach dem „Umgang mit Leistungsunterschieden“ nicht unbedingt der Fall. Sofort möglich wäre eine Einschätzung

des Gelingens zu erbitten, welches auf einer Skala von „1 = sehr gut“ bis „6 = schlecht“ durch Lehrende zu bewerten ist. Aber welche Bedingungen, subjektiven Eindrücke und Erfahrungswerte bei dem Urteil eine Rolle spielen, bleibt unbekannt. Bereits die aufgeworfene Fragerichtung verweist auf ein zentrales Merkmal qualitativer Untersuchungen: Fragen nach dem „Umgang mit“ oder den „Bedingungen von“ sind offen angelegt, eine vorherige Standardisierung erfolgt nicht; eine ganz andere Zielsetzung liegt vor, was enorme Konsequenzen für die Durchführung hat.

Zielsetzung qualitativer Verfahren

Qualitative Verfahren zielen darauf, Phänomene zu verstehen und zu rekonstruieren oder anders ausgedrückt: Fragen nach dem „Was“ und „Wie“ stehen im Zentrum. Damit ist eine vorherige Standardisierung nicht möglich („was“ steht ja noch nicht fest) und es geht darum, herauszufinden, wie „es so und nicht anders gekommen ist“³ – im Kern eine Prozessperspektive. Der gesamte Evaluationsprozess ist offen angelegt, nicht nur um subjektive Deutungen kennenzulernen, sondern auch neue Phänomene in der Praxis, im Alltag der Akteure zu entdecken. Bei aller Offenheit ist ein Ausgangspunkt notwendig: Dies ist die (vorläufige) Fallbestimmung. Ein Fall „[nimmt] sich das Ganze eines in Zeit und Raum gebundenen sozialen Phänomens zum Gegenstand“.⁴ Nicht isolierte Merkmale werden untersucht, sondern ganz explizit der Kontext in seiner Komplexität einbezogen. Jeder Fall ist demnach eingebettet, zum Beispiel Schüler

und Schülerinnen in Familien, diese wiederum in ein Milieu und dieses in eine Region. Eine grundlegende Annahme dabei ist, philosophisch ausgedrückt, „dass man das Ganze aus dem Einzelnen und das Einzelne aus dem Ganzen verstehen müsse“.⁵ Die Befragten sind im Sinne dieser Fallbestimmung Mitglieder einer spezifischen sozialen Position (zum Beispiel Lehrpersonen), die mit Strukturen eines Feldes (zum Beispiel Schulwesen) konfrontiert sind. Ihre Perspektiven entwickeln sich aus biografischen, institutionellen und gesellschaftlichen Erfahrungen, die in der sprachlichen Darstellung eng miteinander verwoben sind. Biografische Erfahrungen können höchst individuell sein, aber bereits institutionell zeichnen sich Deutungen und Regelmäßigkeiten ab, die für eine bestimmte Gruppe stehen. Über diese theoretische Rahmung können Aussagen verallgemeinert werden. Ein etwas plakatives Beispiel aus einem Lehrerinnen-Interview kann das Vorgehen gut illustrieren: „Morgens habe ich dann um acht Uhr begonnen mit dem Unterricht.“ Nehmen wir zur besseren Veranschaulichung an, bei der Institution „Schule“ handelt sich um ein unbekanntes Phänomen, dann verweist die Nennung „acht Uhr“ auf eine Struktur des Feldes. Mit mehreren Interviews im Vergleich kann eine Strukturverallgemeinerung vorgenommen werden: Der Unterricht einer Schule beginnt um acht Uhr. Wird diese Struktur interpretiert und abstrahiert, kann von einer zentralistisch organisierten Zeitstruktur gesprochen werden, was auf einen hohen Formalisierungsgrad hindeutet. In diesem Beispiel-Fall

¹ Vgl. Filsinger, Dieter (2016): Evaluationsforschung und informelles Lernen. In: Harring, Marius / Witte, Matthias D. / Burger, Timo (Hgg.): Handbuch informelles Lernen. Interdisziplinäre und internationale Perspektiven. Weinheim/Basel: Beltz Juventa, S. 768.

² Flick, Uwe / Kardorff, Ernst von / Steinke, Ines (2010): Was ist Qualitative Forschung? Einleitung und Überblick. In: (ders.) (Hrsg.): Qualitative Forschung. Ein Handbuch. 8. Auflage. Reinbek bei Hamburg: Rowohlt Verlag, S. 17.

³ Bude, Heinz (2011): Fallrekonstruktion. In: Bohnsack, Ralf / Marotzki, Winfried / Meuser, Michael (Hgg.): Hauptbegriffe Qualitativer Sozialforschung. 3. Auflage. Opladen & Farmington Hills: Budrich, S. 60.

⁴ Kannonier-Finster, Waltraud / Nigsch, O. / Ziegler, M. (2000): Über die Verknüpfung von theoretischer und empirischer Arbeit in soziologischen Fallstudien. In: Österreichische Zeitschrift für Soziologie, 25/2000, H. 3, S. 6.

⁵ Gadamer, Hans-Georg (1972): Wahrheit und Methode. Grundzüge einer philosophischen Hermeneutik. 3. Auflage. Tübingen: J.C.B. Mohr, S. 275.

werden über die Darstellungsweise Akzeptanz und Routine in der Verarbeitung dieser Zeitstruktur deutlich. Auch diese Bewältigungsstrategie könnte sich als typisch für die Gruppe der Lehrpersonen herausstellen. Ein Kontrast mit einem Schülerinnen-Interview kann aber so ausfallen: „Morgens um acht sitz ich im Unterricht, hab meinen Block vergessen und der Tag beginnt schon ätzend.“ Ohne zu weit zu interpretieren: Ein Deutungsmuster von Routine und Akzeptanz drängt sich nicht gerade auf, aber die gleiche Struktur wird sichtbar. In der Regel sind Befragten solche Wissensbestände nicht bewusst (d. h. reflexiv zugänglich). Es handelt sich um implizites Wissen, welches Menschen das Handeln im (Schul-)Alltag erst ermöglicht. Im Unterricht mit einer unruhigen Klasse kann ein Lehrer nicht Daten sammeln, interpretieren und reflektieren, um dann nach einigen Tagen zu einem Ergebnis einer angemessenen Intervention zu kommen. Qualitative Forschung kann das, denn sie ist von Handlungsdruck entlastet. Dafür benötigt es ein hohes Maß an Reflexivität auf Seiten der Evaluatorsin sowie theoretische Sensibilität auch im Sinne von Abstraktionsfähigkeit und letztlich eine Ton-/Bildaufzeichnung, die transkribiert wird bzw. ein Protokoll – womit ein Fremdverstehen, was im Alltag situativ und meist kaum reflexiv vollführt wird, methodisch kontrolliert vorgenommen werden kann.⁶ Erfüllen die Daten das Prinzip der Alltagsnähe, was gleich vorgestellt wird, erlauben sie ein Verstehen der Perspektiven (1), um dann Prozesse zu erklären (2), woran – wenn relevant – Überlegungen zu Verbesserungen

fast unmittelbar angeschlossen werden können (3). Diese Unmittelbarkeit liegt daran, dass die Ergebnisse empirisch im Alltag derjenigen gründen, deren Praxis gestützt oder verändert werden soll.

Qualitative Datenerhebungen: Alltagsnähe steht im Fokus

Die am meisten eingesetzten Methoden sind Interviews in verschiedensten Variationen, Beobachtungen und Gruppendiskussionen, die häufig durch Dokumentenanalysen ergänzt werden. Offenheit zeigt sich in den Interview- und Diskussionsfragen darin, dass in der Regel lediglich Themen eröffnet werden und nach Erfahrungen damit gefragt wird. Die Befragten können auf diese Weise ihre Erzählungen, Beschreibungen und Argumentationen nach ihren eigenen Relevanzen entfalten. Auch Beobachtungen werden so vorgenommen, möglichst dicht die Prozesse des Unterrichtsgeschehens zu beschreiben. So erhalten die gewonnenen Daten eine deutliche Alltagsnähe – und genau das ist gewollt. In thematisch sowie finanziell, zeitlich und personell begrenzten Evaluationen wird aber meist ein Leitfaden eingesetzt, der Themen aufruft, die auf jeden Fall besprochen/betrachtet werden sollten und der dadurch einen etwas engeren Rahmen setzt. Das Interesse an den subjektiven Deutungen und Erlebnissen der schulischen Akteure kann zunächst irritierend sein oder gar banal wirken. Wird eine Schülerin befragt „Erzähl mir von den letzten Prüfungen“, erfordert die Situation auch Vertrauen und das Wissen darum, dass die folgende Erzählung tatsächlich im Interesse

der Evaluation ist. Fazit: Handlungsbedarfe und Handlungsempfehlungen.

Die beiden Methoden quantitativ und qualitativ untersuchen im Prinzip nicht die gleichen Fragestellungen, wodurch sie auf verschiedenen Ebenen steuerungsrelevante Erkenntnisse liefern. Anhand konkreter und vorgefertigter Konzepte, zum Beispiel zu Lesekompetenzen, können quantitative Häufigkeiten und Zusammenhänge für Schüler und Schülerinnen aufgezeigt werden.⁷ Aber die Fragen danach, wie Bildungsprozesse gelingen und dieses Gelingen unterstützt werden kann, sind qualitativ. Sie sind offen angelegt und ermöglichen, auch neue Formen von Herausforderungen und Gelingensbedingungen zu entdecken. Für das Verwertungsinteresse haben quantitative Daten den Vorteil, Handlungsbedarfe aufzuzeigen (z. B. Steigerung der Lesekompetenz). Qualitative Ergebnisse hingegen sind – durch ihre Gründung im Alltag derjenigen, um die es geht – unmittelbar anschlussfähig an Handlungsempfehlungen. Durch den Einbezug der Kontexte geben sie Auskünfte über Bedingungen, beispielsweise unter welchen Bedingungen Lesekompetenzen bei wem besonders gut entfaltet werden.

Jessica Prigge

Wissenschaftliche Mitarbeiterin im Forschungsverbund „Pädagogik der Kindheit“ des Evangelischen Fröbelseminars und der Universität Kassel sowie Mitglied der Forschungsgruppe ForBES (Bildungs- Evaluations- und Sozialstudien) an der Hochschule für Technik und Wirtschaft des Saarlandes.

⁶ Vgl. z. B. Kruse, Jan (2011): Reader: Einführung in die Qualitative Interviewforschung. Überarbeitete, korrigierte und umfassend ergänzte Version.

Freiburg, S. 19ff.; Przyborski, Aglaja / Wohlrab-Sahr, Monika (2014): Qualitative Sozialforschung: Ein Arbeitsbuch. 4., erweiterte Auflage. München: Oldenbourg, S. 14ff.

⁷ Kruse, Jan (2011): Reader: Einführung in die Qualitative Interviewforschung.

Das Kerngeschäft der Evaluationsstelle

Externe Evaluation der Einzelschulen

Die externe Evaluation der einzelnen Schulen stellt die zentrale Aufgabe der Evaluationsstelle dar, und sie ist dementsprechend zeitintensiv und arbeitsaufwändig.

Nach ersten Erprobungen im Schuljahr 2013/2014 wurden in den folgenden Jahren Richtlinien erarbeitet und Verfahrensweisen festgelegt; im vergangenen Schuljahr schließlich wurde ein mehrjähriges Konzept für die externe Evaluation der Einzelschule erstellt, das folgende wesentliche Merkmale aufweist:

- externe Evaluation aller Schulen im Abstand von sechs Jahren
- Sichtung der internen Evaluation der einzelnen Schule als Ausgangspunkt für die externe Evaluation
- Absprachen bzw. Vereinbarungen zwischen der einzelnen Schule und der Evaluationsstelle hinsichtlich Organisation, Durchführung und Ausmaß der externen Evaluation im Sinne einer partizipativen Evaluation
- Orientierung an den Bereichen des Qualitätsrahmens
- Berücksichtigung festgelegter Kriterien für die Auswahl der Interviewpartner und für die Planung der Unterrichtsbesuche
- Unterrichtsbeobachtungen nach definierten Kriterien für „guten Unterricht“
- Rückmeldung zur internen Evaluation der einzelnen Schule drei Jahre nach Durchführung der externen Evaluation

Für die externe Evaluation einer Schule sind jeweils zwei Mitarbeiterinnen oder Mitarbeiter der Evaluationsstelle als Hauptverantwortliche zuständig; bis zu fünf weitere Personen (je nach Größe und Komplexität einer Schule) arbeiten beim Schulbesuch selber und bei der Erstellung des Rückmeldeberichtes mit. Eine externe Evaluation läuft in folgenden Schritten ab:

- Kontaktaufnahme mit der Schulführungskraft
- Analyse der schulischen Dokumente, im Besonderen zur internen Evaluation
- Planungsgespräch
- Information der Schulgemeinschaft
- Online-Befragung der Lehrpersonen, Schülerinnen / Schüler und Eltern, fallweise zusätzlicher Personengruppen

- Schulbesuch mit Unterrichtsbesuchen sowie Interviews mit Lehrpersonen, Schülerinnen/ Schülern und Eltern
- Auswertung aller erhobenen Daten und eingeholten Informationen und Ausarbeitung des Rückmeldeberichtes
- Besprechung des Rückmeldeberichtes mit der Schulführungskraft
- Präsentation des Rückmeldeberichtes im Lehrerkollegium und auf Wunsch für Eltern und/oder Schülerinnen/Schüler
- Übermittlung des Rückmeldeberichtes an die Schulführungskraft und an den Schulamtsleiter/Ressortdirektor

Einige Zahlen zur externen Evaluation im Schuljahr 2016/2017:

extern evaluierte Schulen	20
durchgeführte Online-Befragungen	69
versendete Fragebögen	20.168
vollständig beantwortete Fragebögen	12.999
durchgeführte Unterrichtsbeobachtungen	566
durchgeführte Interviews	677

Der Schulbesuch

Der Schulbesuch stellt ein zentrales Element der externen Evaluation dar. Wenn ein solcher ansteht, beginnt der Arbeitstag für die Mitarbeiterinnen und Mitarbeiter der Evaluationsstelle meist früh, weil die jeweilige Schule zu Unterrichtsbeginn erreicht werden muss und der Unterricht an vielen Schulen zeitig beginnt. Das Team von Evaluatoren und Evaluatoren ist ausgestattet mit Laptop, Übersichtstabelle mit Einteilung der Klassenbesuche, Zeitplan für die Interviews und dem Bogen für die Unterrichtsbeobachtungen. Mit dem Betreten des Schulgebäudes und ersten Begegnungen mit Lehrpersonen und nichtunterrichtendem Schulpersonal, dann auch mit Schülerinnen und Schülern, beginnt der Schulbesuch, und bereits in diesen Momenten können wesentliche Elemente der vorherrschenden Schulkultur wahrgenommen werden.

Der Vormittag ist ausgefüllt mit Unterrichtsbesuchen. Die Einteilung dafür wird im Vorfeld

so gemacht, dass im Laufe des Vormittags in der Regel alle Schulstellen besucht werden und der Unterricht in möglichst vielen verschiedenen Klassenstufen bzw. Klassenzügen und unterschiedlichen Fächern beobachtet werden kann. Vor Ort müssen allerdings immer wieder Anpassungen vorgenommen werden: An einer Schule verläuft eben nicht jeder Tag gleich und nach Stundenplan. Für die Unterrichtsbeobachtungen wird ein einheitlicher Beobachtungsbogen verwendet, der sich an Kriterien für guten Unterricht orientiert und folgende Bereiche umfasst: Klassenführung, lernförderliches Klima, Motivierung, Klarheit, Wirkungsorientierung, Schülerorientierung, aktives Lernen, Üben, Heterogenität, Variation der Methoden und Sozialformen. Es handelt sich im Übrigen um einen Unterrichtsbeobachtungsbogen in Anlehnung an IQESonline (www.iqesonline.net). Im Laufe des Vormittags werden – mit Ausnahme der Grundschulen – auch Interviews mit Schülerinnen und Schülern geführt. Dafür werden einheitliche Leitfragen verwendet, damit die Gespräche und deren Ergebnisse vergleichbar sind.

Am Nachmittag führen die Evaluatoren Interviews mit einer Reihe von Eltern und Lehrpersonen. Die Gespräche dauern etwa 30 Minuten und werden wie bei den Schülerinnen und Schülern anhand von spezifischen Leitfragen geführt. Der Tag des Schulbesuchs ist für die Evaluatoren und Evaluatoren immer ein intensiver Tag – bei besonders komplexen Schulrealitäten wird der Schulbesuch auch auf zwei Tage ausgedehnt. Am Ende des Tages kann immer wieder festgestellt werden, wie zahl- und facettenreich die Informationen und Eindrücke von einer Schule sind, die das Team im Laufe eines Tages sammeln kann, und wie sich diese mit den Ergebnissen aus den Fragebögen und der Dokumentenanalyse – ähnlich einem Puzzle – zu einem Ganzen zusammenfügen.

Rosa Maria Niedermair

Evaluationsstelle für die deutsche Schule

Qualitätsmanagement an der Berufsschule

Sinn und Zweck von Evaluationen



Pädagogischer Tag an der Landesberufsschule für Handwerk und Industrie in Bozen: Im Rahmen eines Workshops konnten sich Lehrpersonen in ihrer Tätigkeit als Evaluatorinnen und Evaluatoren professionalisieren.

Im Rahmen des Pädagogischen Tages vom 13. März 2017 fand ein Workshop für Lehrerinnen und Lehrer der Berufsschule für Handwerk und Industrie statt. Ziel dieses Workshops war es, die berufliche Evaluationskompetenz der Lehrpersonen zu steigern und sie in ihrer Tätigkeit als Evaluatorinnen und Evaluatoren zu professionalisieren.

Der Workshop gliederte sich in drei Teile. Am Vormittag gab die Direktorin der Evaluationsstelle der deutschen Schule, Ursula Pulyer, eine Einführung zur Evaluation in der autonomen Schule. Schwerpunkte waren dabei die gesetzlichen Ausgangsbedingungen der internen und externen Evaluation sowie eine Erläuterung der Begriffe, die im Evaluationskontext zu finden sind, da es hierzu immer wieder Verwirrungen gibt. Zudem wurde auf die Instrumente eingegangen, die bei der externen Evaluation zum Einsatz kommen. Die Dokumentenanalyse, der Unterrichtsbeobachtungsbogen, die Fragebögen und die Interviewleitfäden orientieren sich jeweils am Qualitätsrahmen, der für alle Schulen verbindlich ist. Dieser beschreibt in den zwei Bereichen Inputqua-

litäten und Prozessqualitäten jene Felder des schulischen Handelns, die maßgeblich das Gelingen von Schule, von Lernen und Lehren und kollegialem Miteinander prägen. Die Wirkungsqualitäten beschreiben die Lernergebnisse der Schülerinnen und Schüler, gemessen an ihrem Bildungserfolg, an den erworbenen Fachkompetenzen und der Persönlichkeitsbildung. Anschließend wurde der Ablauf eines Schulbesuches vom Erstkontakt bis zur Präsentation der Ergebnisse im Detail besprochen. Alle drei involvierten Zielgruppen einer externen Evaluation – Schülerinnen und Schüler, Lehrpersonen, Eltern – erhalten Fragen, die inhaltlich gleich ausgerichtet sind und sich an den Qualitätsrahmen anlehnen. Diese sogenannte Triangulation ist wichtig, um die Aussagen der drei Zielgruppen miteinander vergleichen zu können und die unterschiedlichen Perspektiven klar zu erkennen. Es wurde auch näher auf die Sinnhaftigkeit der partizipativen Evaluation eingegangen, die das Ziel verfolgt, durch ein besseres Zusammenwirken von interner und externer Evaluation Schulentwicklungsprozesse anzuregen.

Feedbackkultur für eine produktive Weiterentwicklung

Im zweiten Teil nahm Alexander Plattner vom Bereich Innovation und Beratung mit den Lehrerinnen und Lehrern die Schweizer Plattform IQESonline genauer unter die Lupe. Dieses Instrument der internen und externen Evaluation sowie der Selbstevaluation garantiert eine gemeinsame Kommunikationsbasis und eine höhere Akzeptanz aller an einer Evaluation Beteiligten. Unterstrichen wurde die Wichtigkeit der Anonymität bei der Durchführung von Evaluationen; dieser Faktor wird manchmal unterschätzt. Die Teilnehmenden haben dann auch einen Schülerfragebogen für das eigene Fach erstellt, und es gab die Möglichkeit, sich über bereits durchgeführte Evaluationen auszutauschen und Optimierungsvorschläge einzubringen.

Am Nachmittag galt das Hauptaugenmerk dem Umgang mit den Evaluationsergebnissen. Dabei wurde die Wichtigkeit von abzuleitenden Maßnahmen nach der Evaluation sowie die Vernetzung zwischen Schul- und Unterrichtsentwicklung und interner Evaluation unterstrichen. Das schulinterne Qualitätsmanagement wird dabei als Schlüsselprozess für die Verarbeitung auch extern gewonnener Monitoring-Informationen zu einer produktiven Weiterentwicklung der Bildungsqualität an den einzelnen Schulen gesehen. Als Basis dieses Prozesses ist eine gut implementierte Feedbackkultur von größter Wichtigkeit, so wie es etwa Bill Gates mit folgenden Worten ausdrückt: „We all need people who will give us feedback. That's how we improve.“

Helga Huber

Evaluationsstelle für die deutsche Schule

Interview mit Rudolf Meraner

Kompetenztests – Chancen und Grenzen

Rudolf Meraner, der Direktor des Bereichs Innovation und Beratung, war mitverantwortlich für die Einführung der Kompetenztests in der deutschen Schule in Südtirol, hat die Verhandlungen mit den deutschen Behörden geführt und vertrat Südtirol lange Zeit in der Steuergruppe von VerA 3 und VerA 6. Er gibt Antwort auf wichtige Fragen, ordnet die Bedeutung der Kompetenztests ein und zeigt deren Grenzen auf.

Was sind Kompetenztests eigentlich?
Rudolf Meraner: Kompetenztests sind Lernstandserhebungen, die zu einem festgelegten Zeitpunkt, in allen betreffenden Klassen zugleich und unter denselben Bedingungen durchgeführt werden, und zwar mit Aufgabenstellungen, die an den zu erreichenden Kompetenzen ausgerichtet sind, wie sie in den Rahmenrichtlinien festgelegt sind. Die Aufgabenstellungen werden in einem mehrjährigen, mehrstufigen, wissenschaftlich begleiteten Verfahren entwickelt, sodass sie möglichst zielgenau das messen, was sie sich zu messen vorgenommen haben.

Sie zeigen also nur einen Ausschnitt?
Rudolf Meraner: Man kann es am besten so erklären: Bewertung von Schülerleistungen vollzieht sich in einem Dreieck von kriteriumsorientierter Norm, individueller Norm und sozialer Norm – das haben alle im Studium und in der Fortbildung gelernt. Die Sozialnorm dürfte inzwischen in der Schule eigentlich keine Rolle mehr spielen und nicht berücksichtigt werden. Die Kunst der Lehrpersonen in der Bewertung ist es also, mit dem Spannungsverhältnis von kriteriumsorientierter und individueller Norm umzugehen. Die Kompetenztests helfen den Lehrpersonen dabei, indem sie ihnen für die kriteriumsorientierte Norm eindeutige, gut abgesicherte Ergebnisse liefern.

Also nur kriteriumsorientierte Norm?
Rudolf Meraner: Ja, das ist wichtig. Ein Kompetenztest zum Lesen in der vierten Klasse Grundschule misst, ob die Kompe-

tenzen in diesem Bereich, wie sie in den Rahmenrichtlinien für das Ende der dritten Klasse angegeben sind, erreicht sind. Diese Aussage müsste für jede Lehrperson von großer Bedeutung sein, denn das Ziel war es ja, den Schülerinnen und Schülern zu helfen, diese Lesefähigkeit zu erreichen.

Und die individuelle Norm?
Rudolf Meraner: Die Kompetenztests können auf diese nicht eingehen. Sie bilden die Leistung zu einem bestimmten Zeitpunkt ab und können keinen Verlauf aufzeigen. Sie können also nicht die Lernfortschritte messen. Deshalb sind die Kompetenztests ein wichtiger Faktor, aber nicht alles. Für eine umfassende Schülerbewertung muss die Lehrperson immer weitere Bewertungselemente heranziehen. Deshalb empfehlen wir auch, die Ergebnisse der Kompetenztests nicht in Noten umzuwandeln, sondern als ein wichtiges Beobachtungselement zu behandeln.

Kompetenztests passen nicht in ein inklusives Bildungssystem, wird manchmal behauptet ...

Rudolf Meraner: Das ist Unsinn. Ich muss aber wissen, was Kompetenztests leisten und was nicht. Sie liefern wichtige Daten, aber diese zu interpretieren, diese anzuwenden, ist Aufgabe der Lehrpersonen. Ich möchte es mit einem Vergleich aus dem Bereich der Medizin erklären: Kompetenztests sind wie ein Blutbild. Die Daten geben dem Arzt grundlegende Informationen, niemand würde darauf verzichten. Aber der Arzt wird sie in einen Zusammenhang stellen mit wei-

teren Daten und Beobachtungen der betreffenden Person, wird sie mit früheren Daten vergleichen. Erst dann kann er den Gesundheitszustand beschreiben. Ein Wert kann für eine Person einen bedeutenden Fortschritt darstellen, für eine andere aber Anlass zu weiteren Maßnahmen sein. Weil es diese Interpretation braucht, kann ich doch nicht sagen, dass das Blutbild sinnlos ist.

Die Daten der Kompetenztests müssen also interpretiert werden?
Rudolf Meraner: Ja, und zwar nur von der Lehrperson der betreffenden Klasse. Diese weiß, was dem Test vorausgegangen ist, welche Maßnahmen sie schon gesetzt hat, wie sich die Klasse im Allgemeinen verhält usw.

Was muss man bei der Interpretation der Daten wissen?
Rudolf Meraner: Gute Kompetenztests sind so aufgebaut, dass der Durchschnitt der Schülerinnen und Schüler eine Lösungshäufigkeit von 55 bis 60 Prozent erreicht. Das ist für viele Lehrpersonen gewöhnungsbedürftig, weil sie ihre Tests in der Klasse danach ausrichten, dass fast alle Schülerinnen und Schüler alle Aufgaben lösen können. Um aussagekräftige Daten für Kompetenztests zu erreichen, müssen die Testaufgaben aber eine viel größere Bandbreite aufweisen. Es gibt bewusst Testaufgaben, die nur von einem Viertel der Schülerinnen und Schüler gelöst werden können, aber auch Testaufgaben, die von 90 Prozent der Schülerinnen und Schüler positiv bearbeitet werden können. Aber erst wenn ich solche Aufga-

ben in einem Kompetenztest habe, kann ich erkennen, ob ich in der Klasse Schülerinnen und Schüler habe, die Kompetenzstufen erreichen, die deutlich über dem Niveau oder deutlich unter dem Niveau des in den Rahmenrichtlinien angegebenen Zielniveaus sind.

Warum braucht es dann Vergleichsdaten?

Rudolf Meraner: Die Vergleiche helfen, die Daten einzuordnen. Bei den Kompetenztests VerA erhalten die Lehrpersonen zum Vergleich einen korrigierten Mittelwert. Dieser sagt: Wie hätte die Klasse abgeschnitten, wenn sie eine durchschnittliche Südtiroler Klasse mit derselben besonderen Zusammensetzung gewesen wäre? Der Vergleich schärft das Bild. Das Entscheidende aber ist, dass ich mit einem forschenden Blick an die Ergebnisse des Kompetenztests herangehe und nicht mit einem strafenden. Was sagen mir die Ergebnisse? Wo liegen die Stärken und Schwächen meiner Klasse? Hätte ich etwas anders machen können oder sollen? Was könnte ich als Nächstes tun? Vor allem die letzte Frage ist eine sehr wichtige!

Welche Aufgabe haben die Schulführungskräfte?

Rudolf Meraner: Für sie gilt genau dasselbe: mit einem forschenden und nicht mit einem strafenden Blick an die Ergebnisse herangehen. Sie sollen nicht einen Vergleich zwischen den Klassen anstellen und dann jene tadeln, deren Klassen schlechtere Ergebnisse erzielt haben. Sie sollen auch den forschenden Blick haben: Welche Hypothesen könnte ich aufstellen, um das Abschneiden

meiner Schule zu erklären? Welche Gründe könnte es für die Unterschiede zwischen den Klassen meiner Schule geben? Und vor allem die Hattie-Frage: Was kann ich als Nächstes tun? Was die Schulführungskraft in jedem Falle tun sollte, ist nachzufragen, ob sich die Lehrpersonen mit den Ergebnissen auseinandergesetzt haben, ob sie für sich Schlussfolgerungen gesetzt haben. Dabei ist auch wichtig, die Lehrpersonen zu bestärken, sich nicht nur mit den Schwächen zu befassen, sondern vor allem die Stärken weiterzuentwickeln.

Und was gibt es noch zu den INVALSI-Tests in der Mittelschule zu sagen?

Rudolf Meraner: Das ist eine Sonderform, weil sie zur Bewertung herangezogen wurden. In den ersten Jahren hat das INVALSI den großen Fehler gemacht, die Mathematik-Tests auch auf eine Lösungshäufigkeit von 55 bis 60 Prozent hin zu kalibrieren, was dann zu unterdurchschnittlichen Noten geführt hat. Die Deutsch-Tests, für die wir in Südtirol Verantwortung übernehmen durften, waren von Anfang an aussagekräftig und auf das Leistungsniveau sehr gut ausgerichtet. Diese sind deshalb von den Lehrpersonen auch besser angenommen worden. Die Kombination von Kompetenztests mit Bewertungssituationen ist aber problematisch und wird deshalb auch wieder zurückgefahren.

Was kann man aus diesen Ergebnissen herausholen?

Rudolf Meraner: Eigentlich dasselbe wie aus den Ergebnissen der Kompetenztests. Das

einzigste ist, dass sich meine forschenden Fragen in die Vergangenheit richten, weil ich diese Schülerinnen und Schüler ja dann nicht mehr unterrichte. Wenn ich bei Mittelschulprüfungen anwesend bin, frage ich mich immer, wie die sogenannte INVALSI-Note mit den übrigen Noten der Mittelschulprüfung übereinstimmt. In den allermeisten Fällen gibt es eine hohe Übereinstimmung. Manchmal kommt es aber vor, dass eine Schülerin oder ein Schüler bei der INVALSI-Note um zwei Stufen über den anderen Noten liegt. Ich beobachte dann immer, dass dies sehr zurückhaltende Schülerinnen oder Schüler sind, die sich nicht so gut ausdrücken können, aber eigentlich alles verstanden haben. Wenn dies häufiger vorkommt, lohnt es sich hinzuschauen und vielleicht diesen Typ von Schülerin oder Schüler unter die Lupe zu nehmen.

Wo gibt es für die Lehrpersonen Unterstützung?

Rudolf Meraner: Neben den Fortbildungsveranstaltungen zum Thema bieten die Fachdidaktikerinnen und Fachdidaktiker am Bereich Innovation und Beratung Unterstützung für Schulen und vor allem für Fachgruppen an. Sie kommen an die Schule oder in die Bezirke und helfen, die Daten zu interpretieren und die richtigen Fragen zu stellen.

Interview:

Redaktion INFO

Der Mythos des guten Indikators

Das Messen des Nicht-Messbaren

Indikatoren sind Anzeiger. Zumindest wenn der Begriff wörtlich übersetzt wird. Aber was damit verbunden ist, bleibt häufig ein für viele unerschlossener Mythos. Warum brauchen wir denn einen Anzeiger? Wie funktioniert das „Anzeigen“ eigentlich? Was ist der Unterschied zum Messen? Und natürlich die Frage aller Frage für jede und jeden, der Indikatoren entwickeln soll: Wie wird ein perfekter Anzeiger gebaut?

Manchmal haben wir Fragen, auf die es keine einfachen Antworten gibt – und die wir trotzdem dringend wissen wollen. Es sind häufig Alltagsfragen wie: Bin ich gesund? Haben meine Schüler und Schülerinnen verstanden, was ich gesagt habe? Bin ich noch auf dem richtigen Weg oder habe ich mich verfahren? Wir wissen genau, dass sich diese Fragen nicht sofort und eindeutig mit „Ja“ oder „Nein“ beantworten lassen – und in der Regel sind sie auch nicht problemlos messbar: Wie misst der Arzt ohne Hilfsmittel und ausreichend Zeit Gesundheit, der Lehrer Lehrqualität oder der Navigator Richtungen? Wenn wir ehrlich sind, heißt die Antwort: gar nicht! Und trotzdem erkennt der erfahrene Arzt, Lehrer, Navigator sofort, wenn etwas nicht stimmt.

Bleiben wir bei dem Beispiel, welches uns selbst am nächsten ist – wieso weiß eine Lehrperson, ob ihr Unterricht bei den Schülern und Schülerinnen ankommt, obwohl sie es nicht gemessen hat? Die Antwort ist: Sie verwendet einen (oder sogar mehrere) Indikatoren. Das können triviale und leicht erkennbare Dinge sein (z. B. wenn der Lärmpegel in der Klasse steigt) oder auch schwierige Einschätzungen (z. B. ob die schwächsten Schüler und Schülerinnen

noch „bei der Sache“ sind oder schon „abgeschaltet“ haben). Dies geschieht so selbstverständlich und alltäglich, dass es uns gar nicht mehr auffällt.

Äpfel mit Birnen vergleichen

Erste Erkenntnis: Indikatoren sind keineswegs komplexe Messinstrumente, die mit unheimlichem Aufwand betrieben werden müssen. Zumeist sind sie sogar das Gegenteil: Die Botschaft muss einfach und leicht verständlich sein – selbst wenn sich dahinter ein kompliziertes Messkonzept verbergen sollte. Nehmen wir wieder ein einfaches alltägliches Beispiel: Wir sagen oft und gerne, dass man Äpfel nicht mit Birnen vergleichen kann – und tun es praktisch bei jedem Marktbesuch. Oder hat schon mal jemand aus Versehen Birnen gekauft, obwohl er oder sie Äpfel haben wollte?

Was tun wir, damit das nicht passiert? Wir vergleichen beispielsweise die Form der Früchte und da diese sich so eindeutig unterscheiden, dass keine Verwechslung möglich ist, gelingt uns der Einkauf unter Verwendung dieses Indikators – der Form – immer und ohne großes Nachdenken. Viel schwieriger ist es, eine bestimmte Apfelsorte zu kaufen – weil sich zwei verschiedene

Sorten in den äußerlich erkennbaren Merkmalen eventuell sehr ähnlich sind und sich in Farbe, Form, Geruch und vielleicht sogar im Geschmack gleichen. Im Korb haben wir dann „Mischobst“...

Wir verwenden also im Alltag bestimmte Eigenschaften oder Dinge, die sich stark voneinander unterscheiden, als Indikatoren, um uns zwischen zwei (oder mehr) Alternativen zu entscheiden. Wenn diese Eigenschaften wie z. B. die Form der Birne immer mit dem entsprechenden Gegenstand verbunden sind und sich immer von dem anderen, auszuschließenden Gegenstand – in diesem Fall der Form des Apfels – unterscheiden, funktioniert das perfekt.

Wir können uns allerdings das Leben sehr schnell schwer machen – nämlich indem wir anfangen richtig zu messen! Plötzlich ist der Apfel unter dem strengen Blick des Maßstabs gar nicht mehr so rund wie er uns vorher erschien und die Grenze, ab wann er „birnig“ wird, verschwimmt. Je genauer wir messen, desto mehr „Streuobst“ erhalten wir – äpfelige Birnen und birnige Äpfel. Es ist also offensichtlich nicht die Genauigkeit der Messung, die uns beim Entscheiden hilft, sondern genau das Gegenteil – ein relativ grobes aber sauber trennendes

Messkonzept. Und da ist es wieder – das Geheimnisvolle des Indikators. Indem er das Nicht-messbare in einfache, leichtverständliche Anzeigen überträgt, macht er uns das Entscheiden leicht. Sobald wir uns um gute Messungen bemühen, wird unsere Entscheidung schwerer ...

Das Leben wäre so einfach, wenn der Indikator aber nicht auch ein Schelm wäre. Die „Messqualität“ besteht nämlich nicht nur aus der Genauigkeit und der Tatsache, dass man aus Metern Millimeter und aus Millimetern Mikrometer machen kann. Fehler entstehen ebenfalls dadurch, dass sich das Nicht-Messbare hinter der „falschen“ Alternative des Anzeigers versteckt. Dies kann der Apfel im Birnenkleid sein – aber auch der kluge Schüler, der mit Hilfe eines Feuerzeugs das Schulthermometer manipuliert, weil er weiß, dass sein Lehrer dies als Indikator für „hitzefrei“ verwendet ...

Der nächste Mythos des Indikators liegt hier begründet: Weil wir das Nicht-Messbare eben nicht messen können, wissen wir auch nicht, auf welche verrückte Ideen die Schülerinnen und Schüler kommen oder in wie vielen Äpfeln der „Birnenwurm“ drin ist. Wenn wir unseren Anzeiger naiv verwenden, dann wird er uns auch immer wieder in die

Irre leiten. Das gilt für den Navigator genauso wie für den Arzt oder die Lehrperson. Was haben wir nun für das Entwickeln von Indikatoren gelernt? Erstens muss uns der Indikator eine klare und eindeutige Handlungsentscheidung geben – zum Beispiel Kaufen oder nicht Kaufen. Zweitens muss der Indikator das Nicht-Messbare exakt dem zuordnen, was wir als Grundlage der Entscheidung verwenden – ein Apfel muss immer rund bleiben und darf nicht birnig werden. Drittens müssen wir mit einer für die Anzeige ausreichenden Genauigkeit messen – wir dürfen für die Entscheidung weder zu ungenau noch zu genau die Form des Obstes untersuchen. Und viertens schließlich müssen wir das Ganze im Alltag jederzeit mit den uns zur Verfügung stehenden Möglichkeiten ausreichend gut umsetzen können. Gerade deshalb bleibt ein guter Indikator leider häufig ein unerreichbarer Mythos. Es kann aber viel Spaß machen, sich dieser Herausforderung zu stellen ...

Wolfgang Meyer

Stellvertretender Leiter des Centrums für Evaluation
CEval der Universität des Saarlandes
Saarbrücken

Die Evaluationsstelle ist für die Organisation und Abwicklung aller Kompetenztests zuständig:

- Schnittstelle zwischen den Zuständigen für die Aufgabenentwicklung bzw. Datenerhebung und den Schulen;
- Übersetzung der Testarbeiten des INVALSI und aller weiteren für die Durchführung der Tests erforderlichen Dokumente;
- Organisation des Drucks und Verteilung der Testunterlagen an die Schulen bzw. Implementierung der Testfragen;
- digitale Abwicklung;
- Auswertung der landesweiten Daten und Erstellung der Landesberichte;
- Ansprechpartner für die Schulen zu sämtlichen Fragen hinsichtlich der Kompetenztests.

Service



Antworten

Frequently Asked Questions

1. Wer legt den Zeitpunkt für die externe Evaluation fest?

Alle Schulen werden in einem Abstand von sechs Jahren evaluiert. Die Auswahl der zu evaluierenden Schulen wird von der Evaluationsstelle festgelegt, mit dem Ziel, jährlich eine Ausgewogenheit hinsichtlich Schultypen, Schulstufen sowie Bezirken zu garantieren.

2. Wie repräsentativ sind die Unterrichtsbesuche?

Das Ziel der Unterrichtsbesuche durch das Evaluationsteam besteht darin, einen breit gefächerten Einblick in das Unterrichtsgeschehen zu erhalten. Dazu wird eine Vielzahl von Stunden nach denselben Kriterien beobachtet (adaptierter Unterrichtsbeobachtungsbogen von IQES-online) und anschließend ausgewertet. Grundsätzlich geht es nicht um die Beurteilung einer einzelnen Lehrperson, sondern um die Einschätzung der Unterrichtsqualität an einer Schule. Die Beobachtungen sind zutreffend, weil Unterricht geprägt ist durch viele routinehafte Abläufe, die sich in den seltensten Fällen kurzfristig herstellen oder vortäuschen lassen.

3. Warum geben die Evaluatoren und Evaluatorinnen kein individuelles Feedback?

Um ein konstruktives Feedback zu einer Unterrichtseinheit zu geben, fehlt den Evaluatoren und Evaluatorinnen Kontextwissen. Sie haben keine Informationen darüber, was in den Stunden vorher unterrichtet wurde, wie die Lehrperson die Stunde geplant hat, welche Ziele sie verfolgt oder welches Vorwissen bei den Lernenden vorhanden ist. Außerdem ist die Beurteilung einer Einzelstunde nicht Zuständigkeit der Evaluationsstelle; deren Aufgabe ist es, der Organisation Schule als Ganzes Rückmeldung zu geben.

4. Wie repräsentativ sind die Interviews?

Bei den verschiedenen Instrumenten, die im Rahmen der externen Evaluation eingesetzt

werden, kommen quantitative und qualitative Aspekte zum Tragen. Die Interviews gehören zur qualitativen Datenerhebung, wobei die Auswahl der Interviewpartner nach ausgewählten Kriterien erfolgt („Fallauswahl“ siehe unten) und daher ein Abbild der Gesamtheit darstellt. Die Mitarbeiter der Evaluationsstelle halten sich bei der Fragestellung an einen Leitfaden und stellen keine Suggestivfragen. Die Interviews werden aufgenommen und können so zum Zweck der verlässlichen Datenauswertung nochmals angehört werden. Ziel der qualitativen Datenauswertung ist es, durch den Vergleich der Aussagen in den verschiedenen Interviews das Gemeinsame, also das, was von allen Befragten so oder ähnlich, formuliert wurde, herauszuarbeiten. Wenn exemplarische Einzelaussagen im Bericht angeführt werden, werden sie auch als solche gekennzeichnet.

5. Wie werden die Interviewpartner ausgewählt?

Die Evaluationsstelle gibt Kriterien vor, nach denen die Fallauswahl erfolgt (z. B. Geschlecht, Dienstalter – nur für Lehrkräfte –, Zugehörigkeit zu einer Schulstelle, Funktion). Bei den Eltern können die Kriterien erst in einem zweiten Moment angewandt werden; zunächst müssen die grundsätzliche Bereitschaft und die zeitliche Verfügbarkeit gegeben sein.

6. Warum wird nicht unterrichtendes Personal nicht befragt?

Die Überlegung, Schulfürer/Schulfürerinnen und Verwaltungspersonal nicht zu befragen, ist eine bewusste Entscheidung, da es bei der externen Evaluation in erster Linie um die Erhebung der Unterrichtsqualität geht. Ein weiteres Kriterium ist die Gewährleistung der Anonymität, die bei einer kleineren Anzahl von Befragten nicht gewährleistet ist.

7. Warum erfolgt die Auswertung nicht nach Schulstellen?

Ziel und Zweck der externen Evaluation ist der Gesamtblick auf die jeweilige Schuldi-

rektion, nicht auf einzelne Klassen, Schulstellen oder Lehrkräfte. Zudem wäre für die vielen kleinen Schulen die Anonymität der Evaluation nicht gewährleistet. Da sich einzelne Schulstufen bzw. Schulzweige innerhalb einer Direktion aber doch sehr voneinander unterscheiden, werden bei Schulsprengeln Grund- und Mittelschule sowie bei Oberschulzentren in der Regel die verschiedenen Schulzweige gesondert betrachtet.

8. Wer erhält den Evaluationsbericht?

Den Evaluationsbericht erhalten nach der Vorstellung im Plenum die Schulführungskraft und der Schulamtsleiter bzw. Ressortdirektor. Der Bericht soll in der Folge allen befragten Mitgliedern der Schulgemeinschaft (Lehrpersonen, Eltern, Schülern und Schülerinnen) vollinhaltlich zur Verfügung gestellt und veröffentlicht werden.

9. Wer gibt Anregung zur Verbesserung der Schulqualität?

Anregungen und Tipps zur Verbesserung der Schulqualität gibt auf Nachfrage durch die Schule der Bereich Innovation und Beratung. Im Aufgabenprofil der Evaluationsstelle ist dies nicht vorgesehen.

10. Wer kontrolliert die Maßnahmen?

Die Schulführungskraft und die Lehrpersonen sind angehalten, die Ergebnisse selbstkritisch zu lesen, auf sie zutreffende Aussagen zu identifizieren und entsprechende Konsequenzen daraus abzuleiten und umzusetzen. Die Mitarbeiter und Mitarbeiterinnen der Evaluationsstelle geben drei Jahre nach der externen Evaluation eine Rückmeldung zum internen Qualitätsmanagement der Schule.

Eva Oberhuber

Evaluationsstelle für die deutsche Schule

Evaluation, Kompetenzerhebungen und Co. zum Nachlesen

Buchtipps und Rezensionen

Sonderheft der „Zeitschrift für Erziehungswissenschaft“, Nr. 16, September 2013:
„Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“



In zwölf Artikeln unterschiedlicher Autorinnen und Autoren kann man das Spannungsfeld, in dem Kompetenzerhebungen sich bewegen, aus verschiedenen Perspektiven betrachten:

einerseits als Diagnoseinstrument mit formativer Absicht zur besseren Förderung und Individualisierung von Lernprozessen, andererseits als summativen Prozess des Bildungsmonitorings. Die optimale Nutzung des Potenzials von formativen und summativen Assessments durch Lehrkräfte zur Weiterentwicklung von Unterricht ist ein von Susanne Bögeholz und Sabina Eggert zentral formulierter Aspekt in ihrem Beitrag zur Rolle von Kompetenzdiagnostik im Rahmen von Lern- und Lehrprozessen. Ebenso spannend klingt z. B. auch die Frage von Oliver Wilhelm bezüglich der Abgrenzung von etablierten Kategorien wie Fähigkeit, Fertigkeit oder Intelligenz.

THOMAS BRÜSEMEISTER, KLAUS-DIETER EUBEL (HRSG.):

Evaluation, Wissen und Nichtwissen
VS-Verlag für Sozialwissenschaften, 2008



„Mit Evaluation sollen Folgen von Handlungen im System sichtbar gemacht werden, um Folgeverantwortung zu stärken und um zukünftiges Handeln an empirisch belegten

Erfahrungen ausrichten zu können.“ Diese plausible und nachvollziehbare Aussage betrachten die Autorinnen und Autoren des Bandes differenziert kritisch und versuchen, die Implikationen, die mit evaluationsbasierten Systemsteuervorhaben einhergehen, aufzuzeigen. Ausgehend von der Perspektive der Educational Governance gehen sie der Annahme nach, dass Evaluation mehr Wissen über das System erzeuge und dadurch eine Steuerung desselben besser möglich wäre, und relativieren sie zum Teil. Interessant ist die weiterführende Auseinandersetzung darüber, ob das vermehrte Wissen auch mit vermehrtem Nichtwissen einhergeht, und welche schulische Akteure welches Wissen benötigen. Im Beitrag von Nils Berkemeier „Schulleitung zwischen Evaluation und Organisation“ wird der Fokus auf die Schnittstelle von Organisation, Schulleitung und Evaluation gerichtet, und es werden die möglichen Ansätze einer Auseinandersetzung von Schulleitung mit Evaluationsergebnissen skizziert. Möglichkeiten und Grenzen unterschiedlicher standardisierter Verfahren werden analysiert und beschrieben.

VERA HUSFELDT

Wirkungen und Wirksamkeit der externen Schulevaluation – Überblick zum Stand der Forschung, in: Zeitschrift für Erziehungswissenschaft Nr. 14, Juni 2011, Ss. 259–282.

Schulinspektionen erlangen in verschiedenen europäischen Ländern einen immer höheren Stellenwert. Gleichzeitig steigen damit auch die Erwartungen in ihre Wirksamkeit, wobei vor allem Wirksamkeit in Hinblick auf nachhaltige Schulentwicklung gemeint ist. Ausgehend von einer Darstellung der Funktionen von Schulinspektionen gibt der Artikel eine Übersicht über Wirksamkeitsmodelle und Forschungsarbeiten. Teilweise wird in den Forschungsarbeiten davon ausgegangen, dass die Wirksamkeit von Inspektionen hinsichtlich der Schulentwicklung direkt an den Schülerleistungen messbar ist. Allerdings bleibt unklar, wie die schulinternen Verarbeitungsprozesse als Bindeglied zwischen Schulinspektion und verbesserter Schülerleistung gestaltet sind. Andere Studien betrachten die Einstellungen oder Reaktionen der Akteure auf die Inspektion, wobei teils Wirksamkeitserwartungen und tatsächliche Schulentwicklung vermischt sind. Ausgehend von der Rekonstruktion der unterschiedlichen Forschungsansätze zeigt der Artikel den Bedarf für Modelle auf, welche die schulischen Verarbeitungsprozesse im Anschluss an Inspektionen stärker in den Blick nehmen.

HOLGER GÄRTNER

Praxis und Verhältnis interner und externer Evaluation im Schulsystem im internationalen Vergleich, in: Zeitschrift für Erziehungswissenschaft Nr. 16, Dezember 2013, Ss. 693–712.

Der Überblicksartikel thematisiert das Zusammenspiel interner und externer Evaluation im Gesamtsystem schulischer Steuerung. Zunächst werden beide Evaluationsansätze skizziert, aktuelle Erkenntnisse über die Wirkung auf die Entwicklung von Schulqualität dargestellt sowie das Verhältnis zwischen normativen Setzungen und Bildungsforschung diskutiert. Beide Evaluationsansätze richten ihr Augenmerk weniger auf schulische Outputs wie z. B. Schülerleistungen, sondern vielmehr auf schulische Prozesse wie Schulmanagement oder die Qualität des Unterrichts. Externe Evaluation in Form von Schulinspektion erfüllt vornehmlich die Funktion von Rechenschaftslegung und der Identifizierung schwacher Schulen. Zur Unterstützung der Schulentwicklung ist die interne Evaluation vermutlich besser geeignet. Um sie jedoch als systematischen Prozess der Qualitätsentwicklung in Verantwortung der Einzelschule betreiben zu können, fehlt es aktuell an der nötigen Kompetenz. Überlegungen zum zukünftigen Verhältnis von interner und externer Evaluation richten sich darauf, externe Evaluation als „Meta-Evaluation“ zu betreiben. Bevor externe Evaluation jedoch in dieser Art auf Ergebnisse interner Evaluation aufbauen kann, ist sowohl ein breiter Kompetenzaufbau in den Schulen als auch ein anderes Rollenverständnis der Inspektorinnen und Inspektoren nötig.

Evaluationsstelle für die deutsche Schule u.a. (HGG.):

PISA 2015 Ergebnisse Südtirol
Bozen 2017



In zehn Kapiteln gehen die Autorinnen und Autoren auf die Ergebnisse Südtirols der PISA-Studie 2015 ein. Nach einer ersten allgemeinen Beschreibung von PISA und der Schilderung der

Durchführung derselben in Südtirol wird ein besonderes Augenmerk auf die ladinische Beteiligung an der Studie geworfen. In einem weiteren Kapitel wird ausführlich auf die naturwissenschaftlichen Kompetenzen der Fünfzehnjährigen eingegangen, wobei der nationale und internationale Vergleich zu den Ergebnissen anderer Länder im Vordergrund steht. Im selben Modus werden in den weiteren Kapiteln die Ergebnisse in Mathematik, Lesen und Financial literacy dargestellt. Den Abschluss bildet eine Trendstudie zu den Veränderungen der PISA-Ergebnisse in Südtirol. Tragendes Element des Berichts ist der Versuch der Verfasserinnen und Verfasser, die Ergebnisse der deutschen, italienischen und ladinischen Schulen in Südtirol in einen nationalen und internationalen Kontext einzuordnen.

Rosa Anna Ferdigg und Ivan Stuppner
Evaluationsstelle für die deutsche Schule

Glossar

BIB

Bereich Innovation und Beratung
(im Deutschen Bildungsressort der
Autonomen Provinz Bozen)
Befasst sich mit den pädagogischen
Aufgaben im Bildungsbereich und verfolgt
das Ziel, die Qualität der Bildungsangebote
zu sichern und weiterzuentwickeln
<http://www.bildung.suedtirol.it/>

Gütekriterien der Evaluation

Folgende drei Kriterien finden Anwendung:
Reliabilität bedeutet die Zuverlässigkeit
einer Messung, d. h. die Angabe, ob ein
Messergebnis bei einem erneuten Versuch
bzw. einer erneuten Befragung unter den
gleichen Umständen stabil ist.

Validität drückt aus, ob ein Messverfahren
oder eine Frage bezüglich ihrer Zielsetzung
geeignet ist.

Objektivität ist dann gegeben, wenn die
Ergebnisse von Evaluationen (z. B. Tests)
unabhängig vom Messenden sind.

IEA

International Association for the Evaluation
of Educational Achievement
Es handelt sich dabei um eine internationale
Organisation, die in einem Zusammen-
schluss mehrerer nationaler und behördli-
cher Forschungsinstitute besteht und sich
in der Bildungsforschung einbringt. Diese
Organisation wurde im Jahr 1958 gegründet.

INVALSI

Istituto nazionale per la valutazione del sis-
tema educativo di istruzione e di formazione
Eine wesentliche Aufgabe des INVALSI ist
die Durchführung der nationalen Lernstand-
serhebungen im Auftrag des Unterrichtsmi-
nisteriums.

<http://www.invalsi.it/invalsi/>

IQB

Institut für Qualitätsentwicklung
im Bildungswesen der Humboldt-
Universität Berlin

Unter anderem entwickelt das Institut
Aufgaben für die Vergleichsarbeiten (VerA),
die regelmäßig in Deutschland und auch in
Südtirol durchgeführt werden.

<https://www.iqb.hu-berlin.de/>

IQESonline

IQESonline ist eine Online-Plattform für
Qualitätsentwicklung und Selbstevaluation
an Schulen. Sie richtet sich an Schulleitun-
gen, Qualitätsbeauftragte, Schulbehörden
und Lehrpersonen. Neben Evaluationsin-
strumenten finden sich auch Fachartikel,
Unterrichtsmaterialien und Beispiele aus
der Praxis.

<https://www.iqesonline.net/>

Lernstandserhebungen, Kompetenztests, VerA

Schriftliche Arbeiten in Form von Tests,
welche die Kompetenzen von Schülerinnen
und Schülern flächendeckend und jahr-
gangsbezogen zu einem bestimmten
Zeitpunkt in verschiedenen Unterrichts-
fächern überprüfen.

Die Bezeichnung der Tests variiert von
Land zu Land, das Akronym VerA steht
für Vergleichsarbeiten.

<https://www.iqb.hu-berlin.de/vera>

Lösungshäufigkeit

Indikator für den Schwierigkeitsgrad
einer Aufgabe

Es gilt: Je höher die Lösungshäufigkeit einer
Aufgabe ist, desto niedriger ist der Schwie-
rigkeitsgrad – oder umgekehrt.

Monitoring

Unter Monitoring versteht man die konti-
nuierliche Erfassung von Daten, um das
(Bildungs-) Management über geplante und
tatsächliche Entwicklungen zu informieren
und abzugleichen, ob festgelegte Standards
über einen längeren Zeitraum hinweg
erreicht werden.

OECD

Organisation for Economic Co-operation and Development
Internationale Wirtschaftsorganisation, die 35 Mitgliedsstaaten umfasst und im Jahre 1960 gegründet wurde. Neben vielen anderen Tätigkeiten im Bezug zum Wohlstand und wirtschaftlichem Fortschritt organisiert die OECD seit dem Jahr 2000 die PISA-Studie.

Pilotierung

Die Aufgaben von Lernstandserhebungen werden im Vorfeld des eigentlichen Haupttests an einer großen Anzahl von Schülerinnen und Schülern erprobt, wodurch die Testeignung und die Schwierigkeit der Aufgaben bestimmt werden.

PIRLS

Progress in International Reading Literacy Study
Internationale Studie, die alle fünf Jahre durchgeführt wird, um die Lesekompetenz der Schüler und Schülerinnen der vierten Schulstufe zu erfassen. Sie wird von der IEA organisiert.

PISA

Programme for International Student Assessment
Internationale triennale Studie zu den Leistungen der Fünfzehnjährigen in Lesen, Mathematik und Naturwissenschaften. Sie wird von der OECD organisiert.

Psychometrie

Zweig der Psychologie, der Fähigkeiten und Verhalten eines Menschen mit empirischen Verfahren erhebt

Qualitätsrahmen

Der Qualitätsrahmen ist ein für die Schulen in Südtirol verbindliches, von der Landesregierung genehmigtes Dokument, das in systematischer Weise die Merkmale und Kriterien guter Schule beschreibt. In den zwei Bereichen Inputqualitäten und Prozessqualitäten beschreibt er jene Felder des schulischen Handelns, die maßgeblich das Gelingen von Schule, von Lernen und Lehren und von kollegialem Miteinander prägen. Die Wirkungsqualitäten beschreiben die Lernergebnisse der Schülerinnen und Schüler, gemessen an ihrem Bildungserfolg und an den erworbenen Kompetenzen.

quantitative – qualitative Datenerhebung

In der Praxis der Sozialwissenschaften unterscheidet man zwei Formen der Datenerhebung. Bei der **quantitativen** Datenerhebung geht es um das zahlenmäßige Erfassen von empirischen Sachverhalten. Hierbei kommen verschiedene Formen der Statistik zum Einsatz (eher deduktiver Zugang). Bei der **qualitativen** Datenerhebung wird über die Methoden der Beobachtung, des Interviews, der Diskussion, der Inhaltsanalyse u. v. m. ein Zugang zu sozialen Phänomenen gesucht. Als Charakteristikum von qualitativen Methoden der Datenerhebung wird oft ihre Offenheit und Flexibilität gegenüber dem Forschungsgegenstand genannt (eher induktiver Zugang).

Testdomäne

getesteter Bereich (z. B. Rechtschreiben oder Leseverständnis im Deutsch-Test)

Testitem

einzelne Testaufgabe

TIMSS

Trends in International Mathematics and Science Study
Internationale im Vierjahresrhythmus stattfindende Studie zu den Leistungen der Schüler und Schülerinnen der vierten und achten Schulstufe in Mathematik und Naturwissenschaften. Sie wird organisiert von der IEA. Es gibt auch eine TIMSS advanced Studie für die Lernenden des Abschlussjahres der Sekundarschule.

Triangulation

Darunter versteht man die Betrachtung eines Sachverhalts aus mehreren Blickwinkeln, z. B. die Beleuchtung eines Themas durch die Anwendung verschiedener Methoden (Beobachtung, Interview, Fragebogen) oder durch den Einbezug verschiedener Personengruppen (z. B. Lehrpersonen, Eltern, Schülerinnen und Schüler).



Evaluationsstelle für die deutsche Schule

Freiheitsstraße 15 | 39100 Bozen

evaluationsstelle@schule.suedtirol.it

www.provinz.bz.it/evaluationsstelle-deutschsprachiges-bildungssystem/



Informationsschrift für Kindergarten und Schule in Südtirol November 2017

Herausgeber: Deutsches Bildungsressort
Presserechtlich verantwortlich: Martin Ebert

Redaktion dieser Ausgabe: Ursula Pulyer
und die Mitarbeiter und Mitarbeiterinnen
der Evaluationsstelle

Fotos: Fotos: Archiv Deutsches Bildungsressort,
Erich Dapunt (S. 30, 36), Gerlinde Egger (S. 5),
Helga Huber (S. 25)

Zeichnungen: Doris Zelger

Anschrift und Sitz der Redaktion:
Deutsches Bildungsressort,
Amba-Alagi-Straße 10, 39100 Bozen,
Tel. 0471 417511, Fax 0471 417519,
redaktioninfo@schule.suedtirol.it
www.provinz.bz.it/schulamt

Grafik: Othmar Schönafinger
Druck: Südtirol Druck, Tschermes

Eingetragen beim Landesgericht Bozen
unter Nr. 18 vom 26.09.2002.
Gedruckt auf chlorfrei gebleichtem,
pH-neutralem Papier, 100% recycelbar.
Verwendung und Nachdruck von Texten nur mit
Angabe der Quelle gestattet.