

Die VergleichsArbeiten VerA 3 Erfahrungen und Resümee nach zehn Jahren

Der Grundschulverband begleitete VerA von Beginn an mit Analysen der Aufgabenqualität, der Durchführungsbedingungen, der Auswertung der Ergebnisse und des Umgangs mit ihnen. Nach einer Diskussion mit dem Schulausschuss der KMK im Jahr 2011 keimte die Hoffnung auf, einen qualitativen Schritt in der Entwicklung von VerA angeregt zu haben. 2012 unterzeichnete die KMK Vereinbarungen zu VerA und traf Verabredungen für Empfehlungen. Ernüchterung brachte die Durchführung 2013, da weder die Verbesserung der Aufgabenqualität eintrat noch Unterstützungsmaßnahmen für Schulen, deren Ergebnisse Bedarfe signalisierten, in Gang gesetzt wurden.

Die Vergleichsarbeiten sind daran zu messen, ob sie den schulpolitischen Anliegen gerecht werden:

- Liefert VerA 3 verlässliche Daten zum Qualitätsstand der Schulen im jeweiligen Land und taugt es deshalb für die Schulentwicklung?
- Liefert VerA 3 verlässliche Daten zur einzelnen Schule und zu den einzelnen Kindern?
- Dient VerA 3 der Implementation der Bildungsstandards?
- Führt VerA 3 zu mehr Qualitätsentwicklung?

Heute muss festgestellt werden, dass eine unter den Bundesländern abgestimmte Zielklarheit bezogen auf VerA schwerlich zu erkennen ist.

Die Vergleichsarbeiten lösen nicht ein, was sie versprechen: ein Instrument für eine verlässliche, objektivierte Leistungsfeststellung des Systems Schule auf seinen verschiedenen Ebenen zu sein. Qualitätssicherung ist unbestritten eine wichtige Aufgabe der Schule. Aber sie ist anspruchsvoll und zugleich schwierig: Die Verfahren dürfen die untersuchten Aktivitäten nicht verfälschen. Notwendig ist ein breites methodisches Repertoire, zu dem u.a. auch Tests zählen. Sie stärken den Außenblick und bieten Vergleichsdaten aus repräsentativen Stichproben. Aber Tests sind auch ein heikles Instrument. Im Rahmen von VerA sind deren Risiken nicht ausreichend bedacht. Schon 1999, lange vor VerA, hat der Grundschulverband mit seinen „5 Thesen zur Evaluation“ eine Konzeption vorgelegt, wie die Rechenschaftsaufgaben im Schulsystem sinnvoll verteilt werden könnten und welche Verfahren dabei pädagogisch angemessen wären.

Warum gibt es Kritik? Was läuft schief?

1. Die Aufgaben erfassen wesentliche Kompetenzen der Kinder nicht.

Im Bereich ‚Texte schreiben‘ beziehen sich z.B. wesentliche Kompetenzen auf den Schreibprozess: Texte planen und Texte überarbeiten. Mit den Aufgabenformaten bei VerA 3 sind solche Kompetenzen nicht erfassbar. Der Grund für die Ausklammerung wesentlicher Kompetenzen liegt in der Eigenart der Testformate und der Auswertung. Die Testaufgaben bei VerA 3 müssen einfach gehalten sein, denn die Lösungen müssen von den Lehrkräften eindeutig und rasch auszuwerten und in eine Tabelle einzugeben sein.

Diese Einschränkungen gelten für alle Bereiche des Deutsch- wie des Mathematikunterrichts. Gerade die komplexeren und deshalb besonders bedeutsamen Kompetenzen sind mit den bei VerA gewählten Testformaten nicht zu ermitteln. Damit sind die Vergleichsarbeiten nicht geeignet,

umfassende oder doch zumindest wesentliche Aussagen zum Kompetenzstand der Kinder zu machen. Für die Qualitätsentwicklung der Schulen sind sie deshalb wenig brauchbar.

2. Die Auswertung bewertet richtige Lösungen der Kinder als falsch.

Ein Beispiel (2008)

Die Kinder sollen in einem Text alle Verben unterstreichen. Wenn Kinder 4 von 6 Verben richtig unterstrichen haben, gilt die Lösung dennoch als falsch – ebenso falsch wie die Lösungen der Kinder, die keines und falsche Wörter unterstrichen haben.

Grund für diese Fehlurteile, die einer möglichen förderdiagnostischen Analyse entgegenstehen, ist wiederum die einfache, der Testmetrik geschuldete Auswertung. Abweichende Leistungen, in denen Kinder ihr Können unter Beweis stellen, werden dabei nicht erfasst – wie auch die entwicklungsbedingte Qualität der Fehler nicht, die oft nur durch eine kontextbezogene Deutung erfasst werden kann.

3. Die Aufgabenstellungen machen leistungsschwächere Kinder mutlos.

Z.B. sollen Drittklässler in dreißig Minuten zwei ganzseitige Texte erlesen und dazu insgesamt 15 Aufgaben bearbeiten (2010). Sowohl langsame als auch besonders sorgfältige Leser sind oft mutlos angesichts dieser Fülle, insbesondere auch Kinder, für die Deutsch die erste Fremdsprache ist.

Die Lesetexte sind häufig so gewählt, dass Kinder sie nicht verstehen, obwohl sie lesen können.

Beispiele

Ein Text berichtete von einem Ideen-Wettbewerb für die Speisekarte des Maritim-Hotels am Potsdamer Platz (2007), ein anderer Text vom Elefantenzug durch ein „Buschhotel“ in Sambia (2010). „Gaststättenbesuch“, „Buschhotel“ – hier ist Weltwissen erforderlich, das viele Kinder in dritten Klassen nicht besitzen. Der Text ist angereichert mit Begriffen, die vielen Kindern unbekannt, aber zum Verständnis zwingend notwendig sind, wie Hotelrezeption, Empfangshalle oder Tresen.

Das Hörbeispiel 2013 erzählte aus der Welt der Wasserwuschel und erforderte bei der Beantwortung von Fragen eine hohe Abstraktions- und Transferleistung, welche für Kinder, deren Muttersprache nicht Deutsch ist, nach einmaligem Zuhören kaum möglich ist. Sie werden regelrecht ausgegrenzt, unbedacht ihrer z.T. hohen Leistungen im Rahmen der Zweisprachigkeit.

Hinzu kommt, dass die Tests gegen Ende der Klasse 3 sich auf Standards beziehen, die in den Bildungsstandards für das Ende von Klasse 4 ausgewiesen sind. Kinder aus bildungsferneren Milieus sind auf schulisches Lernen in besonderer Weise angewiesen. Was sie im Unterricht noch nicht lernen konnten, wissen sie auch noch nicht. Dadurch ergeben sich keine aussagekräftigen Hinweise auf die Potenziale der Kinder und für deren Förderung.

4. Der Umgang mit VerA ist nicht generell geregelt, dem Missbrauch sind Tür und Tor geöffnet.

Die Ergebnisse von VerA 3 werden z.B. bei Schulinspektionen einbezogen. Schulen berichten, dass sie ermahnt wurden, ihre Leistungen bis zum nächsten Jahr zu bessern. Dies geschieht in der Regel ohne Analyse, was unter besseren Ergebnissen für den jeweiligen Standort zu verstehen ist. Es geschieht in der Regel auch ohne Unterstützungsmaßnahmen für die Schule und führt dazu, dass VerA-Aufgaben bereits vorher mit den Kindern geübt werden, dass Kinder bei der Durchführung massive Unterstützung durch die Lehrkraft erhalten oder bei der Eingabe der Ergebnisse nachgeholfen wird.

In einigen Bundesländern werden die Ergebnisse den Eltern mitgeteilt, Schulergebnisse werden öffentlich gemacht, die Ergebnisse des einzelnen Kindes werden als Leistungsnachweis verstanden. Letzteres ist schon testmethodisch ein Missbrauch.

Ein einheitlicher Umgang mit den Ergebnissen ist bisher nicht zustande gekommen, ebenso wenig wie eine faire Vergleichbarkeit, da die erhobenen Daten eine Vergleichbarkeit verschieden situierter Schulen oder gar Klassen nicht sichern.

5. VerA 3 ist unbrauchbar, weil das Konzept des jährlichen Massentests der grundlegende Konstruktionsfehler ist.

Die Vergleichsarbeiten werden jedes Jahr deutschlandweit in allen 3. (und allen 8.) Klassen geschrieben und anschließend ausgewertet. Ein solcher Aufwand ist mit geschultem Personal und mit aufwändigeren Methoden nicht möglich. Deshalb müssen die Lehrkräfte selbst die Organisation, Anleitung und Auswertung übernehmen. Das erzwingt einfache Papier-Bleistift-Tests und einfache Auswertungsschlüssel.

Die jährliche Durchführung mit veränderten Aufgaben erzwingt einfache Aufgabenformate. Nur dadurch wird möglich, Aufgaben mit jeweils anderen Inhalten in kurzer Zeit zu konstruieren, in Vortests zu prüfen und daraus den Test für das jeweilige Jahr zu erstellen.

Diese Bedingungen schließen offenere und didaktisch anspruchsvollere Aufgaben mit diagnostischer Auswertungsqualität aus. Die Tests sind outputorientiert, ohne ausreichend die Ausgangslage der Schülerschaft einer Schule damit in Beziehung zu setzen. Auch daher löst VerA 3 weder die gewünschte Bereitschaft zu interner Evaluation noch zu Qualitätsentwicklung aus.

Die politische Intention, über jährliche Massentests Unterrichtsqualität verbessern zu wollen, erweist sich als grundlegender Konstruktionsfehler.

6. VerA passt nicht zum Auftrag einer auf Inklusion ausgerichteten Schulentwicklung

Seit langem ist den VerA-Verantwortlichen bekannt, dass in Schulen in sozialen Brennpunkten im Schnitt ca. 80 Prozent der Kinder an den Aufgaben scheitern. Dies lässt in vielen Fällen nicht den Schluss zu, diese Schulen seien nicht leistungsfähig oder die Kinder dort alle bildungsfern. VerA versagt hier, denn die Kompetenzen dieser Kinder werden nicht erfasst. Die Vergleichsarbeiten sind auf einen geringen Ausschnitt der schriftlich zu testenden Fachleistungen in Deutsch und Mathematik ausgerichtet und erfassen den Hauptteil der tatsächlich zu erwerbenden Kompetenzen nicht. Vor allem aber sind sie in hohem Maße sprachlastig aufgebaut (auch in Mathematik) und grenzen damit einen großen Teil der Kinder aus. Was geschieht mit lernbehinderten Kindern in der VerA-Situation?

VerA ist in der inklusiven Schule als zentraler Test unangemessen. Schule muss darauf hinarbeiten, die individuellen Leistungen der Kinder zu erfassen und über deren Diagnose die Lernentwicklung zu unterstützen. Vergleichsarbeiten wie VerA sind für eine inklusive Schule ungeeignet. Sie beschämen Kinder und machen sie mutlos. Sie konterkarieren eine Unterrichtsentwicklung, die auf kompetenzorientierte Lernformen zielt und fördern stattdessen lehrplanunabhängiges „teaching to the test“.

GSV, GEW und VBE stellten sich dem Schulausschuss der KMK im Herbst 2011 zur Diskussion über VerA

Die Folgen

Vereinbarungen der KMK zur Weiterentwicklung von VerA gemäß Beschluss vom 08.03.2012

Um die erforderliche Weiterentwicklung von VerA in allen Bundesländern sicherzustellen, wurden folgende Vereinbarungen zur Zielbestimmung von VerA, für die Weiterentwicklung der Aufgabenqualität und die Optimierung der Durchführungsmodalitäten und der Ergebnismrückmeldung in den Ländern beschlossen:

1. Zur Zielbestimmung von VerA in den Ländern:

Zentrale Funktionen sind die Unterrichts- und Schulentwicklung jeder einzelnen Schule und die Vermittlungsfunktion bei der Einführung der fachlichen und fachdidaktischen Konzepte der Bildungsstandards.

In diesem Zusammenhang sehen die Länder ab von

- der Veröffentlichung der VerA Ergebnisse,
- der Identifikationsmöglichkeit einzelner Schulen,
- der Benotung.

Die Länder entscheiden, ob die Schulaufsicht und/oder die Schulinspektion Einsicht in die VerA Ergebnisse auf Schul- und Klassenebene erhalten.

2. Zur Weiterentwicklung der Aufgabenqualität:

- Das IQB entwickelt und pilotiert vermehrt Aufgaben zur Erweiterung des Leistungsspektrums.
- Das IQB erhält regelmäßig Rückmeldung zu aggregierten Daten aus den Ländern als Beitrag zur Qualitätsverbesserung in der Aufgabenentwicklung. Diese Daten werden nicht für einen Ländervergleich genutzt.

3. Zur Optimierung der Durchführungsmodalitäten und der Ergebnismrückmeldung in den Ländern:

- In der dritten Jahrgangsstufe wird VerA in allen Ländern grundsätzlich in allen Grundschulen in einem Fach durchgeführt.
- Es wird nur ein Testheft eingesetzt. Die Erweiterung des Aufgabenspektrums im unteren Leistungsbereich kann von den Ländern optional genutzt werden.
- Für die Ergebnismrückmeldung streben die Länder geeignete Referenzgruppen (sogenannter „Fairer Vergleich“) an.
- Testdurchführung und die Korrekturen erfolgen durch Lehrkräfte (in Hamburg zentral durch das Landesinstitut).

Bestandsaufnahme des GSV zur Umsetzung der Vereinbarungen in den Bundesländern nach der VerA 3 Durchführung im Jahr 2013

...mit dem Ergebnis, dass sich der Grundschulverband im Januar 2014 mit einem Brief an Vorsitz und Mitglieder des Schulausschusses der Ständigen Konferenz der Kultusminister der Länder wandte.

Auszüge aus dem Brief

Sehr geehrte Damen und Herren,

im August 2011 befasste sich ein Ausschuss des Schulausschusses der Kultusministerkonferenz mit den Vergleichsarbeiten VerA 3. Eingeladen zu diesem Sondierungsgespräch waren die Vorsitzenden des Grundschulverbandes GSV, der Gewerkschaft Erziehung und Wissenschaft GEW und des Verbandes Bildung und Erziehung VBE. Das IQB war durch Herrn Prof. Dr. Pant vertreten.

Anlass für das Gespräch waren die Diskussion um die Qualität und Effektivität des Instrumentes sowie die Handhabung und der Umgang mit den Ergebnissen in den Bundesländern. Die eingeladenen Verbände analysieren seit Jahren die Vergleichsarbeiten VerA und kommentieren aus ihren Erfahrungen mit der Umsetzung Stärken und Schwächen des Verfahrens.

Die damalige Leiterin des Schulausschusses, Frau von Ilseman, suchte im Herbst 2011 das Gespräch mit den genannten Verbänden, um einen Austausch über die Kritikpunkte einzuleiten und Perspektiven für VerA zu diskutieren. Einen wesentlichen Teil des Gesprächs nahmen die zusammenfassende Darstellung der Rückmeldungen aus der Schulpraxis und von Eltern, bezogen auf die Aufgabenqualität, die Auswirkungen auf die Unterrichtsentwicklung, die Vergleichbarkeit der Durchführung sowie der Durchführungsbedingungen, ein. Am Ende des konstruktiven Austausches eröffneten sich Perspektiven für die Weiterentwicklung der Aufgabenqualität, für notwendige Absprachen zum Umgang mit Ergebnissen und für die Etablierung von Unterstützungs- und Beratungssystemen. Divergierende Einschätzungen über die Durchführungshäufigkeit, über Wirkungen auf die Unterrichts- und Schulentwicklung sowie über die Kosten-Nutzen-Relation des Instrumentes blieben bestehen. Vonseiten der Schulausschussmitglieder wurde deutlich formuliert, wie sehr den Kultusministerien daran gelegen sei, die Akzeptanz für das Instrument VerA auf Schulebene zu erhöhen.

Als Folge dieses Austausches unterzeichnete die Kultusministerkonferenz am 08.03.2012 Vereinbarungen zur Weiterentwicklung der Vergleichsarbeiten und verabredete Empfehlungen zur verbesserten Nutzung der VerA-Ergebnisse. Für die Durchführung 2013 konnte aus Sicht der Verbände die Umsetzung der Verabredungen erwartet werden.

...

Nach der VerA-Durchführung 2013 nahm der Grundschulverband eine Bestandsaufnahme in den Bundesländern vor. Die Zusammenschau ergab, dass wesentliche Aspekte der Vereinbarungen und Empfehlungen der KMK in den Länderministerien entweder nicht zur Kenntnis genommen oder unterschiedlich interpretiert und umgesetzt wurden.

Beispiele:

- Die Aufgabenqualität wies erneut deutliche Mängel auf. Die testtheoretische Auswertbarkeit bestimmte die Systematik der Auswertung und vernachlässigte richtige Teilergebnisse. Lese- und Hörverständnistexte waren unangemessen.
- Die Kriterien für den so genannten fairen Vergleich wurden nach wie vor bundeslandspezifisch getroffen und waren z.T. unscharf bzw. wurden als Schätzwert abgegeben, da den Schulen hierzu Datenangaben fehlen.
- Eine standortbezogene Interpretation der Ergebnisse sowie die Bearbeitung der Konsequenzen hätten Unterstützungsangebote für Schulen vorausgesetzt, die weithin fehlen.

Der Grundschulverband stellt fest, dass die getroffenen Verabredungen auch in der VerA-Phase 2013 keinen zureichenden Niederschlag gefunden haben. Die kooperative Haltung des Verbandes konnte nicht zu der erforderlichen Verbesserung des Instrumentes beitragen.

Ende der Briefzitate

Maresi Lassek, Vorsitzende des Grundschulverbandes
Prof. Dr. Hans Brügelmann Fachreferent „Qualitätsentwicklung“ im Grundschulverband