

## Pädagogische Nutzung der Vergleichsarbeiten

### Vergleichsarbeiten – Chancen und Grenzen

Vergleichsarbeiten dienen der Standortbestimmung schulischer Leistungen, auch im Hinblick auf bundeslandübergreifende Bildungsstandards. Gleichzeitig sollen Vergleichsarbeiten einen Beitrag zur Unterrichtsentwicklung leisten. Vergleichsarbeiten bewirken keine direkte und automatische Verbesserung des Lehrens und Lernens, sondern stellen ein pädagogisch nutzbares Informationsangebot dar. Diese Handreichung möchte das pädagogische Potenzial der Vergleichsarbeiten verdeutlichen.

Vergleichsarbeiten bieten vielfältige Vergleichs- und Interpretationsmöglichkeiten. Ihre Aussagekraft unterliegt aber der Einschränkung, dass es sich nur um eine *Momentaufnahme* handelt, die durch zusätzliche Informationen ergänzt und abgesichert werden muss. Verglichen mit den Erfahrungen und Erkenntnissen von Lehrerinnen und Lehrern kann das Ergebnis einer Vergleichsarbeit im Hinblick auf die Beurteilung einzelner Schülerinnen und Schüler immer nur ein ergänzender Mosaikstein sein – nicht mehr, aber auch nicht weniger.

Wichtig: Vergleichsarbeiten sind nur **ein** Element des Qualitätssicherungssystems. Die konsequente und effiziente Nutzung der Vergleichsarbeiten für die Unterrichtsentwicklung erfordert funktionierende schulische und externe Unterstützungssysteme. Vor diesem Hintergrund sehen wir für die pädagogische Nutzung von Vergleichsarbeiten fünf aussichtsreiche Felder:

#### 1. Beurteilung der Leistungsverteilung der eigenen Klasse

Vergleichsarbeiten ermöglichen einen "fremden Blick" auf das Leistungsprofil der eigenen Klasse durch Vergleichsperspektiven, die im Schulalltag normalerweise nicht zur Verfügung stehen. Dies kann Grundlage für gezielte unterrichtliche Schwerpunktsetzungen sein. Drei Vergleichsebenen sind hervorzuheben:

- *Innerschulische Profile*: Wo hat meine Klasse Stärken und Schwächen, verglichen mit den Parallelklassen? In welchen Kompetenzbereichen zeigen sich Schwächen, die im Hinblick auf die Erreichung der Standards Anlass für gezielte Unterrichts- und Fördermaßnahmen sein müssen?
- *Vergleich mit allen Schülerinnen und Schülern der Jahrgangsstufe*: Wo weicht meine Klasse (nach oben oder unten) von der Leistungsverteilung des eigenen Bundeslandes ab?
- *Fairer Vergleich*: Weicht das Ergebnis meiner Klasse vom Profil einer Gruppe von vergleichbaren Klassen (ähnliches Einzugsgebiet, ähnliche Schülerzusammensetzung) ab?

#### 2. Analyse einzelner Aufgaben

Jenseits der globalen Abschätzung des Leistungsstandes der Klasse, ihrer Stärken und Schwächen, kann auch der Blick auf einzelne Aufgaben didaktische Impulse liefern: Bei welchen Aufgaben ist meine Klasse ausgesprochen erfolgreich, verglichen mit den durchschnittlichen Lösungshäufigkeiten, und woran könnte dies liegen?

Aus fachdidaktischer Sicht kann es aufschlussreich sein, das Gegenstück zur Aufgabenlösung zu betrachten, nämlich die Verteilung von Falschlösungen in der eigenen Klasse.

- *Auffällige Fehlermuster*: Bei welchen Aufgaben zeigen sich in meiner Klasse markante Abweichungen vom Durchschnitt, was die Häufigkeit bestimmter Falschlösungen anbelangt?
- *Zufällige oder systematische Fehler*: Ob Fehler systematisch oder zufällig auftreten, lässt sich auf Grund einer einzigen Vergleichsarbeit mit einer begrenzten Menge an Aufgaben nicht zuverlässig feststellen.

Wir empfehlen: Überlegen Sie in der Fachgruppe zunächst, welche Denk-, Rechen- oder Konzeptfehler etc. hinter besonders häufig auftretenden Fehlern stehen könnten. Grenzen Sie anschließend mögliche Ursachen ein, indem Sie Aufgabenstellungen systematisch variieren. Verwenden Sie hierfür strukturell ähnliche Aufgaben, bei denen jeweils ein schwierigkeitsbestimmendes Merkmal verändert wird, z.B. bei Mathematikaufgaben die Anzahl erforderlicher Lösungsschritte. Auf diese Weise lässt sich nach und nach herausfinden, was eigentlich hinter bestimmten Fehlern steht.

### 3. Von der Diagnose zur Erklärung

Bisher ging es in Lernstandserhebungen nicht um Erklärungen. Aus gutem Grund, denn Vergleichsarbeiten liefern lediglich den "Rohstoff" für Fragen nach den Ursachen für auffällige, „erwartungswidrige“ Klassen- oder Schulergebnisse. Hier liegt ein fruchtbares Feld für die pädagogische Diskussion. Angenommen, eine Klasse weist bei der Vergleichsarbeit in einem Inhaltsbereich deutlich schwächere Leistungen auf als der Durchschnitt (der Schule oder des Bundeslandes). Liegt dies ...

- *an den Schüleinnen und Schülern?* (mangelnde Vorkenntnisse? geringe Anstrengung, z.B. weil es keine Note gab? mangelndes Verständnis der Testinstruktion, z.B. häufige Mehrfachankreuzungen oder Auslassen von Aufgaben anstatt intelligenter Schätzung?)
- *am Unterricht?* (Wurde der Stoff noch nicht oder seit langem nicht mehr behandelt? Erfordern die Aufgaben Kompetenzen, die im Unterricht keine nennenswerte Rolle spielten? Waren spezifische Fähigkeiten – wie z.B. das Schätzen (in Mathematik) oder die Analyse diskontinuierlicher Texte (in Deutsch) – möglicherweise (noch) nicht Gegenstand des Unterrichts?)
- *am Lehrbuch?* (Ist das Lehrbuch "nahe dran" oder "weit weg" von Kompetenzen, wie sie in den Vergleichsarbeiten, die sich an den Standards orientieren, erwartet werden? Begünstigt es bestimmte Fehlerarten?)
- *an den Testaufgaben?* (Weicht Ihre persönliche Einschätzung der "Richtigkeit" einer bestimmten Aufgabe von den Vorgaben der Korrekturanweisung ab?)

Je nach Schwerpunkt der Diskussion sind unterschiedliche Konsequenzen denkbar: gezielte schulinterne Weiterbildung; Initiativen zur Anreicherung des Aufgabenpools (etliche Schulen haben dies bereits getan); Planung von Unterrichtsprojekten, z.B. Einsatz der Videografie zur Verbesserung einzelner Lehr- und Unterrichtskompetenzen; Praktizierung des "lauten Denkens" von Schülern im Unterricht, um bestimmten Fehlern (z.B. im Bereich arithmetischer Operationen in Mathematik oder im Bereich der Orthographie in Deutsch) auf die Spur zu kommen; gezielte Förderprogramme; wechselseitige Hospitation.

### 4. Vergleichsarbeiten als Ausgangspunkt für Förderung

Die Ergebnisse der Vergleichsarbeit können als Startpunkt betrachtet werden, um bestimmte Schülergruppen systematisch zu fördern, spezifische Teilgebiete (Algorithmen, Strategien) gezielt zu üben und vertiefen oder Wissensdefizite ausgleichen. Wiederholt man die Vergleichsarbeit nach einiger Zeit, dann eröffnet sich auf diese Weise die Chance, den Ertrag der Förderung zu prüfen:

- Wer hat von der Förderung profitiert, wer nicht?
- Falls verschiedene Methoden der Förderung oder des Trainings erprobt wurden: Welche hat sich bewährt?

An die Stelle bloßer Vermutungen und Hoffnungen treten hier empirisch begründete Daten zu nachweislichen Effekten gezielter Förderung. In einem solchen wirkungsorientierten Ansatz – Koppelung der Vergleichsarbeit mit gezielter (und selbst evaluierter!) Förderung - liegt nach unserer Einschätzung das pädagogisch größte Potenzial von Vergleichsarbeiten.

## 5. Erfassung der Diagnosegenauigkeit

Bei einem Teil der Aufgaben werden die Lehrpersonen gefragt: "Was schätzen Sie, wie viele Schülerinnen und Schüler Ihrer Klasse diese Aufgabe lösen werden?" (nicht in NRW). Diese Prognosen lassen sich mit den Daten zum tatsächlichen Abschneiden der Klasse bei den entsprechenden Aufgaben in Beziehung setzen:

- *Unter- oder Überschätzung*: Neigen Sie insgesamt (d.h. über alle ausgewählte Aufgaben hinweg) eher zur Unter- oder Überschätzung des Leistungsvermögens Ihrer Klasse? Dies ermöglicht Hinweise darauf, wie realitätsangemessen Ihre Leistungserwartung ist. Wenn sich deutliche Abweichungen zeigen: Wie sind diese zu erklären?
- *Didaktische Sensibilität*: Das Ergebnis der Vergleichsarbeit in Ihrer Klasse kann man nach Aufgabenschwierigkeit sortieren: von der schwierigsten (von wenigen gelöst) zur leichtesten (von vielen gelöst). Nach dem gleichen Prinzip lassen sich Ihre eigenen Einschätzungen ordnen. Aus dem Vergleich beider Rangreihen lässt sich ein Indikator Ihrer "didaktischen Sensibilität" berechnen. Je ähnlicher sich reale und geschätzte Aufgabenschwierigkeiten sind, desto besser sind Sie darüber im Bilde, was Aufgaben in Ihrer Klasse leicht oder schwer macht.

Kennwerte für die Tendenz zur Unter- oder Überschätzung des Leistungsniveaus Ihrer Klasse sowie für die Genauigkeit Ihrer Diagnosen sind Teil der Ergebnismeldung im Internet. Ausführlichere Informationen hierzu finden Sie in der „Handreichung zur Diagnosegenauigkeit“.

## Weiterführende Literatur

- Groß Ophoff, J., Koch, U., Hosenfeld, I. & Helmke, A. (2006). Vergleichsarbeiten in Grundschulen - und was diese daraus machen (können). *Journal für Schulentwicklung*, 10 (4), 7-12.
- Helmke, A. (2004). Vergleichsarbeiten, Parallel- und Orientierungsarbeiten in der Grundschule: Möglichkeiten ihrer pädagogischen Nutzung. *Seminar - Lehrerbildung und Schule*, Heft 2/2004, Themenheft "Lernen und Leisten in der Grundschule", S.90-112
- Helmke, A. & Hosenfeld, I. (2005). Leistungsmessung in der Grundschule: Vorbereitung und Nutzung von Vergleichs- und Orientierungsarbeiten. *Klexer*, 14, 3-5.
- Helmke, Hosenfeld & Schrader (2004). Vergleichsarbeiten als Werkzeug für die Verbesserung der diagnostischen Kompetenz von Lehrkräften. In R. Arnold & C. Griese (Hrsg.), *Schulleitung und Schulentwicklung* (S. 119-144). Hohengehren: Schneider.
- Helmke, A. (2007). *Unterrichtsqualität: Erfassen, Bewerten, Verbessern* (4. Auflage). Seelze: Kallmeyersche Verlagsbuchhandlung.
- Helmke, A. & Hosenfeld, I. (2003). Vergleichsarbeiten (VERA): Eine Standortbestimmung zur Sicherung schulischer Kompetenzen – Teil 2: Nutzung für Qualitätssicherung und Verbesserung der Unterrichtsqualität. *Schulverwaltung*, Ausgabe NRW (5), 143-145.
- Hosenfeld, I. (2005). Rezeption - Reflexion - Aktion: Wie lassen sich Lernstandserhebungen und Vergleichsarbeiten pädagogisch nutzen? In G. Becker, A. Bremerich-Vos, M. Demmer, K. Maag Merki, B. Priebe, K. Schwippert, L. Stäudel & K. J. Tillmann (Hrsg.), *Standards. Unterrichten zwischen Kompetenzen, zentralen Prüfungen und Vergleichsarbeiten* (Friedrich Jahresheft XXIII 2005, S. 112-114). Seelze: Friedrich Verlag.
- Hosenfeld, I., Groß Ophoff, J. & Bittins, P. (2006). Vergleichsarbeiten und Schulentwicklung. *Schulmanagement Handbuch*, 25.
- Hosenfeld, I., Schrader, F.-W. & Helmke, T. (2007). Von der Rezeption zur Ergebnisrückmeldung. Leistungsevaluation im Spannungsfeld von System-Monitoring und Schulentwicklung. In I. Hosenfeld & F.-W. Schrader (Hrsg.), *Schulische Leistungen. Grundlagen, Bedingungen, Perspektiven* (S. 289-313). Münster: Waxmann.
- Koch, U., Groß Ophoff, J., Hosenfeld, I. & Helmke, A. (2006). Qualitätssicherung: Von der Evaluation zur Schul- und Unterrichtsentwicklung - Ergebnisse der Lehrerbefragungen zur Auseinandersetzung mit den VERA-Rückmeldungen. In F. Eder, A. Gastager & F. Hofmann (Hrsg.), *Qualität durch Standards? Beiträge zur 67. AEPF-Tagung (Arbeitsgruppe der Empirischen Bildungsforschung) in Salzburg* (S. 187-199). Münster: Waxmann.