

*Innovativer
Bildungsservice*

Vergleichsarbeiten Baden-Württemberg

**Bericht zur Evaluation der
Vergleichsarbeiten 2009**

Stuttgart 2009



Landesinstitut
für Schulentwicklung

www.ls-bw.de
best@ls.kv.bwl.de

Qualitätsentwicklung
und Evaluation

Schulentwicklung
und empirische
Bildungsforschung

Bildungspläne

Redaktionelle Bearbeitung

Redaktion: Peter B. Sikorski
Autoren: Friedrun Kemmner
Peter B. Sikorski
Sonja Wagner
Stand: Dezember 2009

Impressum

Landesinstitut für Schulentwicklung
Referat 31 „Empirische Verfahren“
Rotebühlstraße 131
70197 Stuttgart

Tel.: 0711/6642-202
Fax: 0711/6642-204

E-Mail: dva@ls.kv.bwl.de
Internet: www.dva-bw.de

© Landesinstitut für Schulentwicklung, Stuttgart 2009

Inhalt

| | |
|--|----|
| Vorbemerkung | 3 |
| 1 Telefonisch und schriftlich am LS eingegangene Rückmeldungen | 4 |
| 1.1 Zur Konzeption | 5 |
| 1.2 Zur Organisation | 6 |
| 1.3 Zur Informationstechnik | 7 |
| 2 Online-Fragebogen | 9 |
| 2.1 Schwierigkeiten organisatorisch-technischer Art | 9 |
| 2.2 Nutzen der Vergleichsarbeiten | 9 |
| 2.3 Passung der einzelnen Vergleichsarbeit | 11 |
| 2.4 Anmerkungen und Verbesserungsvorschläge | 12 |
| 3 Prognostizierte und tatsächliche Lösungshäufigkeiten | 13 |

Vorbemerkung

Der vorliegende Bericht zur Evaluation der Vergleichsarbeiten 2009 gliedert sich in drei Teile.

Kapitel 1 beinhaltet die Zusammenfassung der Rückmeldungen, die telefonisch oder schriftlich als E-Mail, Brief und Fax im Landesinstitut eingingen.

Kapitel 2 präsentiert die Ergebnisse der Auswertung eines Online-Fragebogens, der von den Lehrkräften alternativ zu den telefonischen und schriftlichen Rückmeldungen (Kap. 1) verwendet werden konnte.

Kapitel 3 fasst die Ergebnisse des Vergleichs der Pilotierung 2008 und des Pflichteinsatzes 2009 in Bezug auf die prognostizierten und die tatsächlich Lösungshäufigkeiten zusammen.

Der Erfassungszeitraum reicht vom 18. September 2009 (Bereitstellungstermin der Testmaterialien) bis zum 30. November 2009.

Die Einsatzbedingungen der Vergleichsarbeiten haben sich im Jahr 2009 gegenüber den Vorjahren verändert. Zum einen wurde der Einsatzzeitpunkt der Arbeiten vom Ende des jeweiligen Bildungsabschnitts (z.B. Ende Klasse 8) auf den Anfang des nächst folgenden Bildungsabschnitts (z.B. Anfang Klasse 9) verlegt, zum anderen entfiel die Notengebung.

Darüber hinaus wurde den Schulen neben der Auswertungsmappe „Klasse“ erstmals eine Auswertungsmappe „Schule“ angeboten, mit der automatisierte Parallelklassenvergleiche möglich sind.

Trotz dieser Änderungen blieb die Anzahl der Anfragen und Anmerkungen in etwa gleich. Waren es im Jahr 2008 in einem vergleichbaren Zeitraum insgesamt 170 Meldungen, so stieg die Zahl im Jahr 2009 nur geringfügig auf 195 Meldungen.¹

¹ Nach dem genannten Erfassungszeitraum gingen und gehen weitere schriftlichen Stellungnahmen ein - insbesondere von ganzen Fachschaften - die sich auf einzelne Arbeiten oder allgemein auf die Konzeption der DVA beziehen. Diese „Nachträge“ werden im LS aufmerksam registriert, liefern aber keine neuen Erkenntnisse.

1 Telefonisch und schriftlich am LS eingegangene Rückmeldungen

Beim Landesinstitut für Schulentwicklung (LS) gingen telefonisch oder schriftlich insgesamt 195 Rückmeldungen ein, die sich auf die Kommunikationswege wie in Tabelle 1 dargestellt verteilen.

Tabelle 1: Kommunikationswege zum Landesinstitut

| Eingang | absolute Häufigkeit | prozentuale Häufigkeiten |
|-------------|---------------------|--------------------------|
| per Telefon | 119 | 61% |
| per E-Mail | 70 | 36% |
| per Brief | 6 | 3% |
| insgesamt | 195 | 100% |

Deutlich mehr als die Hälfte der Rückfragen und Rückmeldungen wurden telefonisch übermittelt. Dies zeigt, dass die telefonische Hotline, die den Schulen angeboten wurde, das zentrale Kommunikationsmittel darstellt.

Bei den einzelnen Kontakten wurden häufig mehrere Aspekte angesprochen, wodurch sich die Zahl der im Folgenden aufgeführten einzelnen Rückmeldungen auf 262 erhöht.

Tabelle 2: Betreff der Rückfragen und Rückmeldungen an das Landesinstitut

| Betreff | absolute Häufigkeiten | prozentuale Häufigkeiten |
|----------------------|-----------------------|--------------------------|
| Konzeption | 64 | 24% |
| Organisation | 84 | 32% |
| Informationstechnik | 114 | 44% |
| davon einzelne Tests | 10 | 4% |
| insgesamt | 262 | 100% |

Nicht ganz die Hälfte, und damit die häufigsten Kontakte, betrafen den Bereich der Informationstechnik, rund ein Drittel den Bereich der Organisation. Der Bereich der Konzeption war mit einem Anteil von etwa einem Viertel vertreten. Speziell zu einzelnen Tests gingen weniger als 5% der Rückfragen und Rückmeldungen ein, wobei außer einem Hauptschultest (HS6D² 1-mal) nur Gymnasialtests angesprochen wurden (GY6D, GY6M, GY8M und GY8F je 1-mal, GY8L 2-mal und GY8D 3-mal).

² Die einzelnen Arbeiten der DVA werden im Folgenden so bezeichnet, wobei die ersten beiden Stellen die Schular angeben, die dritte Stelle die Klassenstufe und die vierte Stelle das Fach.

1.1 Fragen/Anmerkungen zur Konzeption der DVA

Zu den einzelnen Bereichen gingen vielfältige Rückfragen und Rückmeldungen ein, die im Folgenden detailliert aufgeführt werden. Die Anzahl der Nennungen ist jeweils in Klammern angefügt.

- Benotung:
 - Frage nach der Erlaubtheit der Benotung der DVA (8)
 - Forderung nach Wiedereinführung der Benotung (7)
 - Zustimmung zum Wegfall der Notengebung (1)
- Motivation der Schüler/innen:
 - Bericht über gute (1) bzw. schlechte (9) Motivation der Schüler/innen
- Zeitlicher und materieller Aufwand:
 - Bericht über zu hohen Zeitaufwand für die Lehrkräfte und Wunsch nach zeitlicher Entlastung (9)
 - Bericht über zu hohen Kopieraufwand (5)
 - Wunsch nach zentraler Auswertung (2)
- Zur Verfügung gestelltes Material:
 - Positive Bewertung der neuen Handreichung zum Umgang mit den DVA (1)
 - Positive Bewertung der neuen Schulmappe (1)
- Verschiedenes:
 - Frage nach der verpflichtenden Durchführung der DVA (8)
 - Anzweifeln der Vergleichbarkeit, da viele Kollegen *teaching to the test* betreiben (4)
 - Kritik am Einsatzzeitpunkt zu Beginn des Schuljahres (4)
 - Verfahrensweise bei Nichtbehandlung verschiedener Inhalte im Unterricht (1)
 - Frage nach den Zielen der DVA (2)
 - Frage nach der Konstruktion der Leistungsgruppen (1)

Die Benotung der DVA ist seit diesem Durchgang nicht mehr relevant. Hierauf beziehen sich 16 Rückmeldungen. 8-mal wurde nachgefragt, ob selbständig benotet werden dürfe, 7-mal wurde geäußert, dass es besser wäre, die DVA wieder zu benoten. Dies wurde begründet mit mangelnder Schülermotivation sowie der zusätzlichen Korrekturleistung, die zu erbringen sei. Nur eine Rückmeldung bewertete die Nichtbenotung als positiv.

Zur Schülermotivation gingen 10 Rückmeldungen ein, wobei die Lehrkraft einer Hauptschulmädchenklasse von einer sehr guten Motivation berichtete, alle anderen jedoch von einer sehr schlechten Motivationslage, vor allem in RS und GY Klasse 9.

Es gab auch wieder einige Rückmeldungen, die den Aufwand an materiellen Ressourcen (Papierumfang und Kopierkosten) kritisierten. Zwei Rückmeldungen forderten aus diesem Grund eine zentrale Auswertung.

Außerdem wurden Anfragen gestellt, ob und in welchem Umfang die DVA verpflichtend seien. Dies war für einige Schulen offensichtlich unklar. Ein Schulleiter wollte wissen, ob die elektronische Dateneingabe verpflichtend sei, da eine Fachschaft diese nicht vornehmen wolle.

Einige wenige hielten die DVA für ungeeignet, Klassen und Schulen miteinander zu vergleichen. Es wurde befürchtet, dass ein „teaching to the test“ die Vergleichbarkeit verfälschte. Weiterhin beklagten einige die Verschiebung des Einsatzzeitpunktes auf den Anfang des darauffolgenden Schuljahres, da ein Lehrerwechsel erfolgt sei. Die Chance, eine Rückmeldung über die Erreichung des vorangegangenen Bildungsstands zu bekommen, wurde anscheinend an dieser Stelle nicht gesehen.

1.2 Zur Organisation

Zur Organisation gingen ebenfalls viele Anmerkungen und Fragen ein. Diese reichten über Anfragen zur Zusendung von Materialien bis hin zu terminlichen Fragen.

- Material:
 - Wunsch nach Zusendung bestimmter Materialien. (20)
- Umgang mit Ergebnissen der DVA intern:
 - Frage nach der Umsetzung von Erkenntnissen aus den DVA (11)
- Umgang mit Ergebnissen der DVA nach außen:
 - Frage nach Art der Rückmeldung der Ergebnisse an Eltern und Schüler/innen? (21)
 - Frage nach der Weitergabe von Testmaterialien an Schüler/Eltern und Frage nach Aufbewahrungspflicht (16)
- Wochentage und Termine:
 - Allgemeine Fragen zu Terminen (7)
 - Fragen zu Terminverschiebungen, Nachschreiben der DVA (5)
 - Wunsch, nicht die gleichen Wochentage für die verschiedenen DVA zu wählen. (2)
- Frage danach, welche Fremdsprache getestet werden soll (2)

Eine größere Personengruppe äußerte den Wunsch, Mappen oder Materialien zugesandt zu bekommen. Meist waren es Vertreter von Schulen in Freier Trägerschaft, die dann darauf hingewiesen wurden, dass es bei den Staatlichen Schulämtern Ansprechpartner gibt, welche die DVA an private Schulen weitergeben, wenn ein Berechtigungsformular vorgelegt wird.

Fast 50 Meldungen betrafen die Thematik, wie mit den DVA-Ergebnissen umzugehen sei: Ob sie im Unterricht und mit Eltern besprochen werden müssten, ob sie nach Hause gegeben werden dürften und ob es eine Aufbewahrungsfrist gäbe. Auch wie die DVA schulintern zu nutzen seien, war Gegenstand der Anfragen.

Das Landesinstitut verwies in diesem Zusammenhang auf die neue Verwaltungsvorschrift vom 17. Mai 2009, die regelt, dass die Vergleichsarbeiten mit Eltern, Schülern und in Lehrerkonferenzen besprochen werden müssen und auf Wunsch mit nach Hause gegeben werden müssen.

Bezüglich der Aufbewahrung wurde empfohlen, mit den DVA wie mit Klassenarbeiten umzugehen, jedoch wurde auch darauf hingewiesen, dass eine ausdrückliche Aufbewahrungspflicht nicht besteht.

Einige wenige Anfragen betrafen die Durchführungstermine. Auch nach Nachschreibterminen wurde gefragt. Den Schulen stand es frei, die DVA nachschreiben zu lassen, sie wurden aber darauf hingewiesen, dass der verschobene Einsatzzeitpunkt einen Einfluss auf die Vergleichbarkeit mit anderen Ergebnissen haben kann.

Zwei Schulen wünschten, dass für die verschiedenen Fächer künftig verschiedene DVA-Wochentage gewählt werden, da sonst in bestimmten Fächern zu viel Unterricht ausfalle.

1.3 Zur Informationstechnik

Bezogen auf die Informationstechnik gingen die meisten Fragen ein, die sich auf das Mitarbeiterportal, die jeweilige Hard-/Software und die Auswertungsmappen bezogen.

- Mitarbeiterportal:
 - Bericht über Probleme mit dem Mitarbeiterportal (21)
 - Bericht über Probleme beim Herunterladen der Mappen (5)
- Hardware/Software:
 - Bericht über Probleme bei der Verwendung von OpenOffice, anderen Betriebssystemen oder anderen Programmen (9)
 - Bericht über Probleme bei der Verwendung von Excel 2000 (32)
 - Bericht über Probleme mit maßstabsgetreuem Drucken/Druckereinstellungen (4)
- Auswertungsmappen des LS:
 - Probleme bei der Bearbeitung der Excelmappe allgemein (4)
 - Nachfrage zum Reparaturprogramm (sog. „Patch“, s.u.) (15)
 - Nachfragen zu Fehlermeldungen in den Mappen (8)
 - Nachfrage zum Ausdruck von Einzelrückmeldungen (4)
 - Frage nach der Möglichkeit des Entfernens doppelt eingelesener Klassen aus der Schulmappe (3)
 - Frage zu angeblich fehlenden Items (2)
 - Bericht über Probleme beim Handling der Schulmappe (1)
 - Fragen zur Dateneingabe und Datenrückmeldung (2)
 - Aussagen zur Qualität der Auswertungsmappe (2)
- Materialien:
 - Wunsch nach Aufhebens des Kopierschutzes im Lehrer-Heft (1)

Unmittelbar nach der Durchführung mussten die DVA-Auswertungsmappen für Benutzer von älteren Excel-Versionen (vor Excel 2000) angepasst werden, da bei ihrer Verwendung ein Schreibschutz auf den Mappen lag und nichts eingetragen werden konnte. Obwohl die Anpassung seitens des LS sofort vorgenommen wurde und die Schulleitungen zeitnah informiert wurden, führte mangelnder Informationsfluss in den Schulen dazu, dass über 30 Anfragen hierzu im Landesinstitut eingingen.

Rund 20 Schulen hatten Probleme mit dem im Sommer 2009 neu eingerichteten Mitarbeiterportal der Kultusverwaltung oder dem Intranet und wurden an das Service-Center Schulverwaltung verwiesen.

15 Schulen bezogen sich auf einen kleinen Programmierfehler in den Auswertungsmappen, der bei einer besonderen Konstellation dazu führte, dass ein bestimmter Schüler trotz richtiger Berechnung individueller Testwerte einer falschen Leistungsgruppe zugeordnet wurde. Den Schulen wurde umgehend ein Reparaturprogramm (sog. „Patch“) zugesandt, mit dem die betreffenden Mappen korrigiert werden konnten. Dies wurde von einem Anrufer ausdrücklich sehr gelobt.

Obwohl den Schulen mitgeteilt wurde, dass zur Bearbeitung der Auswertungsmappen wegen im Hintergrund arbeitenden Makros nur Excel ab Version 2000 verwendet werden darf, gab es einige Anfragen zu mit OpenOffice oder mit anderen Betriebssystemen bzw. anderer Software bearbeiteten Mappen.

Weitere Schwierigkeiten entstanden, wenn Lehrkräfte mittels *Drag and Drop* Teile verschoben. Dadurch wurden bei der Auswertung falsche Bezüge hergestellt.

Nur sehr wenige Schulen hatten ein generelles Problem mit dem Herunterladen der Mappen und mit maßstabsgerechtem Drucken, oder generell mit der Handhabung der Mappe.

Vier Schulen waren von einem weiteren kleinen Programmierfehler betroffen, der nur in einer seltenen Konstellation auftauchte (nur dann, wenn ein Schüler kein Item im ersten Schwerpunktbereich gelöst hatte) und keinen Einzelausdruck von Schülerergebnissen erlaubte (ein Seriena Ausdruck war dagegen immer möglich).

Drei Schulleitungen hatten Klassenmappen versehentlich in die Schulmappe doppelt eingelesen und wollten sie wieder entfernen, was jedoch nicht möglich war. In der im nächsten Jahr zur Verfügung gestellten Mappe wird dies möglich sein.

Bei Tests mit vielen Items konnten in den Excel-Mappen die unteren Items auf dem Tabellenblatt "Eingabe Items" nicht auf nur einer Seite dargestellt werden. Zwei Schulen scrollten nicht weit genug herunter und meinten, es fehlten Items.

2 Online-Fragebogen

An der Befragung nahmen 66 Lehrkräfte teil (HS: 15=22,7%; RS: 19=28,8%; GY: 32=48,5%). Einzelne Lehrkräfte waren für mehr als ein Fach zuständig, in die Antworten gehen daher Erfahrungen mit insgesamt 75 Einsätzen von Vergleichsarbeiten ein.

Diese Datenbasis ist für eine differenzierte Analyse zu klein. Die Antworten der Lehrkräfte sind allerdings relativ stringent, sodass sich doch einige interessante Tendenzen ableiten lassen.

2.1 Schwierigkeiten organisatorisch-technischer Art

Im ersten Frageblock wurde nach Schwierigkeiten organisatorisch-technischer Art gefragt. 54 Lehrkräfte (81,8%) verneinten die entsprechende Frage, 12 (18,2%) bejahten sie. Tabelle 1 zeigt, wo die Schwierigkeiten im Einzelnen festgemacht wurden (insgesamt 14 Antworten).

Tabelle 1: Schwierigkeiten organisatorisch-technischer Art

| Aspekte | Nennungen |
|---|-----------|
| Informations- und Materialienweitergabe an der Schule | 0 |
| Formale Qualität der Schülerhefte (z.B. Layout, Grafiken) | 3 |
| Bearbeitung der Excel-Auswertungsmappe | 7 |
| Rückmeldung der Ergebnisse an Schüler/innen und Eltern | 2 |
| Sonstiges | 2 |

Von den 7 Lehrkräften, die mit der Bearbeitung der Excel-Auswertungsmappe Schwierigkeiten hatten, gehörten 5 der Realschule an.

2.2 Nutzen der Vergleichsarbeiten

Im zweiten Fragenblock waren Aussagen in Bezug auf den Nutzen der Vergleichsarbeiten einzuschätzen (1 = trifft gar nicht zu; 2 = trifft eher nicht zu; 3 = trifft eher zu; 4 = trifft voll und ganz zu).

Tabelle 2 gibt einen Überblick über die gemittelten Einstufungen. Werte kleiner 2,5 können als tendenzielle Ablehnung, Werte größer 2,5 als tendenzielle Zustimmung gelesen werden.

Tabelle 2: Nutzen der Vergleichsarbeiten, gemittelte Einstufungen

| | HS | RS | GY | Gesamt |
|---|------|------|------|--------|
| Ich fühle mich bezüglich der Zielsetzungen der Vergleichsarbeiten ausreichend informiert. | 3,47 | 2,95 | 3,06 | 3,12 |
| Die in der Auswertungsmappe "Klasse" ausgewiesenen Ergebnisse sind für mich verständlich. | 3,67 | 2,95 | 3,53 | 3,39 |
| Ich weiß, wie ich Erkenntnisse aus den Ergebnissen der Vergleichsarbeiten gewinnen kann. | 3,27 | 3,00 | 3,13 | 3,12 |
| Ich setze die aus den Vergleichsarbeiten gewonnenen Erkenntnisse in meinem Unterricht um. | 3,27 | 2,58 | 2,94 | 2,91 |
| Der veränderte Einsatzzeitpunkt begünstigt die Nutzung der gewonnenen Erkenntnisse. | 2,80 | 2,32 | 2,56 | 2,55 |
| Die Vergleichsarbeiten sind für mich als diagnostisches Instrument zum Schuljahresbeginn nützlich. | 2,93 | 2,16 | 2,44 | 2,47 |
| Für die Analyse und Interpretation der Ergebnisse habe ich die "Handreichung zum Umgang mit den Ergebnissen" genutzt. | 2,93 | 2,68 | 3,00 | 2,89 |
| Die Vergleichsarbeiten werden an meiner Schule systematisch analysiert. | 2,80 | 2,00 | 2,66 | 2,50 |
| Aus den gewonnenen Erkenntnissen werden an meiner Schule Konsequenzen für die Weiterentwicklung der Schule gezogen. | 3,00 | 1,89 | 2,53 | 2,45 |
| Insgesamt | 3,13 | 2,50 | 2,87 | 2,82 |

Unter anderem lassen sich folgende Antworttendenzen ablesen:

- Über alle Aspekte und Schularten gemittelt wird der Nutzen der Vergleichsarbeiten tendenziell positiv eingestuft (2,82). Dies gilt vor allem für die Hauptschule (3,13) und auch für das Gymnasium (2,87), weniger für die Realschule (2,50).
- In Bezug auf den veränderten Einsatzzeitpunkt und den damit verbundenen Nutzen der Vergleichsarbeiten ist man geteilter Meinung.
- In der Realschule werden die Ergebnisse der Vergleichsarbeiten nur unzureichend zur Weiterentwicklung der Schule herangezogen.

2.3 Passung der einzelnen Vergleichsarbeit

Beim dritten Fragenblock wurden die Aussagen zur Passung der Vergleichsarbeiten für jede einzelne Arbeit separat erfasst.

Die insgesamt 69 Stellungnahmen verteilen sich auf die einzelnen Vergleichsarbeiten wie folgt:

HS6D = 11; HS6M = 6; RS6D = 5; RS6M = 3; RS8D = 5; RS8M = 3; RS8E = 5; RS8F = 0; GY6D = 8; GY6M = 5; GY8D = 8; GY8M = 3; GY8E = 3; GY8F = 2; GY8L = 2

Da die Fallzahlen für eine Analyse der einzelnen Vergleichsarbeiten zu gering sind, werden die Einstufungen im Folgenden für die Klassenstufen der einzelnen Schulararten zusammengezogen und gemittelt, und zwar in Bezug auf die Fächer Deutsch und Mathematik, für die die meisten Aussagen vorliegen.

1 = trifft gar nicht zu; 2 = trifft eher nicht zu; 3 = trifft eher zu; 4 = trifft voll und ganz zu

Wieder können Werte kleiner 2,5 als tendenzielle Ablehnung und Werte größer 2,5 als tendenzielle Zustimmung gelesen werden.

Tabelle 3: Passung der Vergleichsarbeiten, gemittelte Einstufungen

| | HS6D HS6M | RS6D RS6M | RS8D RS8M | GY6D GY6M | GY8D GY8M |
|--|--------------|--------------|--------------|--------------|--------------|
| Die meisten Schüler/innen waren motiviert, diese Vergleichsarbeit zu bearbeiten. | 2,88 | 2,63 | 1,88 | 2,62 | 2,09 |
| Die Inhalte dieser Vergleichsarbeit passen zu meiner Unterrichtspraxis. | 2,82 | 2,13 | 2,50 | 3,15 | 3,82 |
| Das Anforderungsniveau dieser Vergleichsarbeit war zu hoch | 2,35 | 2,00 | 2,00 | 2,08 | 1,91 |
| Die Bearbeitungszeit war für fast alle Schüler/innen ausreichend. | 2,47 | 3,25 | 3,25 | 3,38 | 2,45 |
| Die Korrekturanweisungen (Lösungen) sind für die objektive Beurteilung der Schülerleistungen geeignet. | 3,29 | 3,25 | 2,63 | 2,54 | 2,82 |

Unter anderem lassen sich folgende Antworttendenzen ablesen:

- Die Motivation der Schüler zur Bearbeitung der Vergleichsarbeiten ist am Anfang von Klasse 9 geringer als am Anfang von Klasse 7.
- Die inhaltliche Passung der Vergleichsarbeiten mit der Unterrichtspraxis wird in der Realschule in Frage gestellt.
- Das Anforderungsniveau der Vergleichsarbeiten war generell angemessen.
- Die Bearbeitungszeit war insgesamt für die meisten Schüler/innen ausreichend (Ausnahme: HS6D, GY8D).
- Die Eignung der Korrekturanweisungen für die objektive Beurteilung der Schülerleistungen wird insgesamt positiv beurteilt (Ausnahme: GY6D).

2.4 Anmerkungen und Verbesserungsvorschläge

Die Lehrkräfte hatten die Möglichkeit, über die standardisierten Fragen hinaus Anmerkungen und Verbesserungsvorschläge für die einzelnen Tests in einem dafür vorgesehenen Antwortfeld frei zu formulieren. Insgesamt machten 21 Lehrkräfte davon Gebrauch (HS = 2; RS = 8; GY = 11).

Da auch hier die Datenbasis für eine inhaltliche Analyse nicht ausreicht, werden im Folgenden die Anmerkungen nur schlagwortartig aufgelistet (in Klammern ist die Anzahl der Nennungen aufgeführt). Zu einem großen Teil decken sich die Inhalte mit den im vorigen Abschnitt beschriebenen Rückmeldungen.

- Forderung nach der Erlaubtheit der Benotung (1)
- Zustimmung zum Wegfall der Benotung (1)
- Zustimmung zum veränderten Einsatzzeitpunkt (1)
- Bericht über zu hohen Kopieraufwand, zu hohen Papierverbrauch (5)
- Bericht über schlechte Motivation von Schülern und Lehrern (3)
- Bericht über zu hohen Zeitaufwand, zu wenig Entlastung (6)
- Bericht darüber, dass die DVA hilfreich bei Schülereinschätzungen sei (1)
- Wunsch nach einem Verfahren zur Onlineauswertung (2)
- Wunsch nach Abschaffung der DVA (4)
- Unkenntnis der Verwaltungsvorschrift in Bezug auf Rückmeldung an Eltern (1)
- Kritik an formaler Qualität der Schülerhefte (1)

3 Prognostizierte und tatsächliche Lösungshäufigkeiten

Wie für die Vergleichsarbeiten der Vorjahre³ wurde auch für die einzelnen Vergleichsarbeiten 2009 untersucht, inwieweit prognostizierte und tatsächliche Lösungshäufigkeiten der Items (Teilaufgaben) übereinstimmen.

Bei den „prognostizierten Lösungshäufigkeiten“ handelt es sich um die in den jeweiligen Auswertungsmappen angegebenen landesweiten Vergleichswerte. Diese wurden auf der Basis der 2008 durchgeführten Pilotierung geschätzt.

Die „tatsächlichen Lösungshäufigkeiten“ wurden dagegen im Rahmen des Pflichteinsatzes der Vergleichsarbeiten am Anfang des Schuljahres 2009/2010 bei einer repräsentativen Stichprobe von speziell dafür ausgewählten Schulen erhoben.

In der nachfolgenden Abbildung ist für die Vergleichsarbeiten 2009 der Hauptschule, der Realschule und des Gymnasiums dargestellt, wie groß im jeweiligen Test der Anteil von Items mit mehr oder weniger großen Differenzen in den Lösungshäufigkeiten ist. Dabei werden fünf „Differenzkategorien“ (DK) unterschieden:

- DK -2:
prozentualer Anteil der Items, die im Pflichteinsatz **deutlich schwerer** ausfallen („Ausreißer“), d.h. von deutlich weniger Schülerinnen und Schülern gelöst werden (Differenz größer -20%)
- DK -1:
prozentualer Anteil der Items, die im Pflichteinsatz **schwerer** ausfallen (Differenz zwischen -11% und -20%)
- DK 0:
prozentualer Anteil der Items, die im Pflichteinsatz der Pilotierung **vergleichbar** ausfallen (Differenz zwischen -10% und +10%)
- DK +1:
prozentualer Anteil der Items, die im Pflichteinsatz **leichter** ausfallen (Differenz zwischen +11% und 20%)
- DK +2:
prozentualer Anteil der Items, die im Pflichteinsatz **deutlich leichter** ausfallen („Ausreißer“), d.h. von deutlich mehr Schülerinnen und Schülern gelöst werden (Differenz größer +20%)

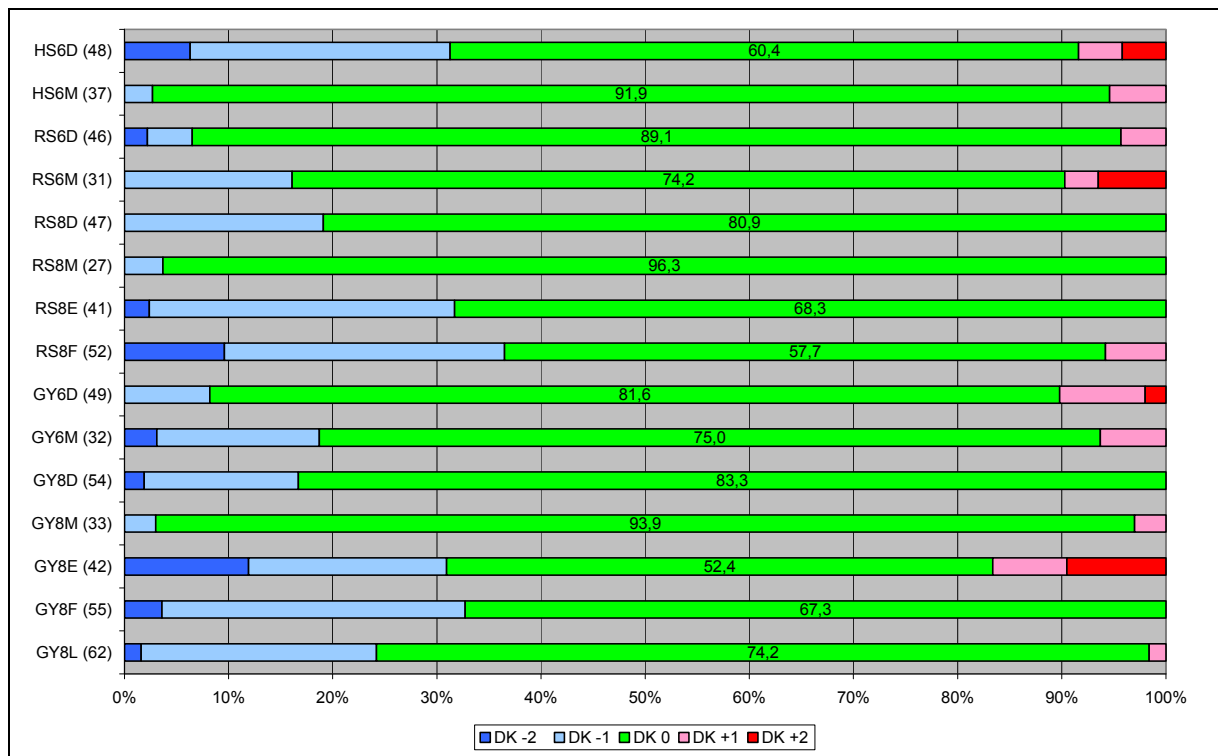
In der Graphik ist für die Differenzkategorie „DK 0“ (grün) angegeben, wie viel Prozent der Items dort jeweils angesiedelt sind. Zur Relativierung der Angaben ist außerdem neben der Kurzbezeichnung der jeweiligen Vergleichsarbeit⁴ in Klammern angegeben, aus wie vielen Items sich die Arbeit zusammensetzt.

Die Vergleichsarbeit „GY8M“ beispielsweise setzt sich aus 33 Items zusammen, 31 (93,9%) liegen im „grünen Bereich“, jeweils 1 Item hat sich im Pflichteinsatz als schwerer (blau) bzw. leichter (rot) erwiesen als prognostiziert.

³ Siehe: www.schule-bw.de/entwicklung/dva/vadva/weiterfuehrend/

⁴ Schulart, Bildungsstandard, Fach. Beispiel: RS8M = Realschule, Bildungsstandard 8, Mathematik

Abbildung 1:
Vergleichsarbeiten 2009, Verteilung der Differenzkategorien (DK)



Es zeigt sich, dass die meisten Items im „grünen Bereich“ (DK 0) liegen, d.h. die Differenz zwischen der prognostizierten Lösungshäufigkeit (Pilotierung 2008) und der tatsächlichen Lösungshäufigkeit (Pflichteinsatz 2009) ist in der Regel kleiner gleich 10 Prozent.

Größere Differenzen zwischen prognostizierten und tatsächlichen Lösungshäufigkeiten treten vor allem bei den Arbeiten auf, bei denen die Itemdefinitionen und die Lösungsvorgaben etwas offener gestaltet waren (z.B. GY8E), oder bei denen die Bearbeitungszeit relativ knapp bemessen war (z.B. HS6D).

Über alle Vergleichsarbeiten gesehen ist der Anteil der Items im „blauen Bereich“ (DK -2 bis DK -1) etwas größer als im „roten Bereich“ (DK 1 bis DK 2), d.h. die Vergleichsarbeiten 2009 fallen im Pflichteinsatz tendenziell eher schwerer aus als aufgrund der Pilotierungsergebnisse angenommen.

Es wird erwartet, dass sich prognostizierte und tatsächliche Werte mit zunehmender Etablierung der 2009 erstmals vorgegebenen neuen Durchführungsbedingungen (Einsatz der DVA am Anfang des dem Bildungsabschnitt folgenden Schuljahrs, Wegfall der Noten) weiter anpassen werden.