

Test

Lernkontrollen als Instrumente der Leistungsbeurteilung und der Diagnose im Geschichtsunterricht

Ob Multiple-Choice-Aufgaben, Lückentexte oder Freiantwort-Aufgaben – auch im Geschichtsunterricht verfehlen Tests ihre Wirkung nicht. Sinnvoll konzipiert und eingesetzt, können sie der Leistungsmessung und Diagnose dienen.

Von Klaus Fieberg

Definition und Gütekriterien

Tests sind ihrer Definition zufolge Verfahren zur quantitativen Bestimmung des relativen Grades von individuellen Merkmalsausprägungen. Einfacher formuliert: Tests messen den Grad des Erreichens von Kenntnissen und Fähigkeiten. Sie unterliegen den Gütekriterien

- a) der Validität beziehungsweise Gültigkeit (Wie genau misst der Test die Kompetenz, die er messen soll? Misst das Verfahren tatsächlich das gewünschte Merkmal bzw. ist es für die Messung des Merkmals tauglich?),
- b) der Reliabilität beziehungsweise Zuverlässigkeit (Wird das Merkmal zuverlässig gemessen oder ist die Messung in zu großem Ausmaß mit Messfehlern behaftet?) und
- c) der Objektivität (Sind die Ergebnisse unabhängig von Einflüssen der Untersucher oder der Untersuchungssituation bei Durchführung, Auswertung und Interpretation zustande gekommen? Wie objektiv wird der Test durchgeführt? Wie objektiv wird der Test ausgewertet?).

Funktionen

Tests dienen auf den ersten Blick zwar vornehmlich der Kontrolle des Lernerfolgs und der Leistungsbeurteilung. Sie können daneben aber auch für eine Beurteilung des Lernzuwachses von Nutzen sein. Sie sind ein wesentliches Instrument der pädagogischen Diagnostik und können als Diagnoseinstrument zur Grundlage für Entscheidungen im Sinne einer individuellen Förderung werden. Tests können außerdem dazu eingesetzt werden, dem Unterrichtenden Rückmel-

dung über die Wirksamkeit seines Unterrichts zu vermitteln. Im Falle der Auswertung eines Tests durch die Schülerinnen und Schüler selbst kann dieser der Selbsteinschätzung der eigenen Leistung und der Motivation für eine Leistungssteigerung dienen.

Standardisierte und informelle Tests

Standardisierte Tests sind im schulischen Rahmen in Form etwa von Intelligenztests, Schuleingangstests, Tests zur Erfassung von Sprachstand und Lesekompetenz, Rechtschreib- oder Rechentests üblich. Im Gegensatz zu standardisierten Tests, die wissenschaftlich entwickelt und unter standardisierten Bedingungen durchgeführt werden, werden informelle Tests durch Lehrerinnen und Lehrer selbst erstellt und von ihnen auf die jeweilige Lerngruppe abgestimmt. Sie orientieren sich an einer speziellen Schülergruppe und an exakt definierten Lerninhalten und Lernzielen.

Testformate

Informelle Tests lassen sich einteilen in solche mit freier oder offener und solche mit gebundener oder geschlossener Aufgabenbeantwortung.

Testformate mit geschlossener Aufgabenbeantwortung

Bei Testformaten mit geschlossener Aufgabenbeantwortung sind die Antwortmöglichkeiten vom Testkonstrukteur bereits vorgegeben. Der Proband hat die Aufgabe, aus vorgegebenen Alternativen die richtige Antwort auszuwählen oder zu konstruieren. Dabei ist der Schüler von

fast aller Schreibarbeit entbunden. Die Auswertung der Ergebnisse dieser Testformen durch den Lehrer erfordert einen relativ geringen Korrekturaufwand, und auch ihre Beurteilung ist in starkem Maße rationalisiert und objektiviert.

Zuordnungsaufgaben

Diese bieten zwei Listen an, deren Elemente jeweils einander zugeordnet werden müssen. Zuordnungsaufgaben fordern etwa dazu auf, sinnvolle Beziehungen zwischen vorgegebenen Namen, Daten, Begriffen oder Fakten herzustellen. Solche Aufgaben sind allerdings nur dann sinnvoll, wenn zwei Kategorien einander eindeutig zugeordnet werden können, etwa Jahreszahlen - Ereignisse; Ereignis - Epoche. Denkbar ist auch, dass beispielsweise Fragen und die passenden Antworten, Begriffe und ihre Definitionen oder Namen und Bilder von Personen einander zugeordnet werden müssen. Ein Problem ergibt sich, wenn Reizwörter und Assoziationswörter in gleicher Anzahl vorhanden sind. Dann nämlich zieht eine falsche Zuordnung automatisch eine zweite nach sich. Um dies zu verhindern, sollte man mehr Assoziations- als Reizwörter vorgeben.

Umordnungsaufgaben

Hierbei müssen einzelne, ungeordnete Begriffe oder Aussagen nach festgelegten Kriterien in eine bestimmte Reihenfolge gebracht werden. Dies kann beispielsweise die korrekte zeitliche Abfolge geschichtlicher Ereignisse sein.

Zweifachwahlaufgaben oder Alternativ-Antworten

Bei diesem Format werden pro Aufgabe genau zwei Antwortalternativen angeboten, eine richtige und eine falsche. Durch

Ankreuzen werden einzelne Aussagen vom Probanden als richtig oder falsch bewertet. Dabei besteht allerdings auch ohne jegliches Wissen aufgrund der Raterwahrscheinlichkeit für den Probanden statistisch eine Chance von 50%, die Aufgaben richtig zu lösen. Dies schränkt trotz der kurzen Lösungszeiten die Tauglichkeit von Alternativ-Antworten (True-false-Items) erheblich ein.

Mehrfachwahlaufgaben oder Multiple-Choice-Aufgaben (MC)

Eine einfache Multiple-Choice-Aufgabe besteht aus einem Aufgabenstamm und mehreren, oftmals vier Antwortoptionen (Alternativen). Eine (oder, bei komplexeren MC-Aufgaben, mehrere) der Antwortoptionen ist (sind) richtig. Der Aufgabenstamm ist so konstruiert, dass ein Problem formuliert wird, zu dem der Proband die richtige Auswahl der Antwortalternativen treffen muss. Für die Formulierung des Aufgabenstamms unterscheidet man im Wesentlichen zwei Formulierungsvarianten:

- a) Aufgabenstamm als Frage: Alternativen als Antworten auf die Frage;
- b) Aufgabenstamm als unvollständiger Satz, der durch jede Alternative vervollständigt wird.

Die falschen Alternativen heißen Distraktoren. Diese zu formulieren, stellt die eigentliche Schwierigkeit bei der Konstruktion brauchbarer Multiple-Choice-Aufgaben dar. Von ihnen hängen Schwierigkeit und Güte der Testaufgabe ab. Der Proband soll allein mit seinem Sachverstand die richtigen und falschen Antworten unterscheiden können. Alle angebotenen Distraktoren sollten im Zusammenhang mit der Aufgabenstellung sinnvoll, logisch und gleich wahrscheinlich erscheinen. Die Antwortvarianten sollten zudem eine vergleichbare Länge und Komplexität aufweisen.

Testformate mit offener oder freier Aufgabenbeantwortung

Im Gegensatz zu der ersten Hauptgruppe müssen die Schüler bei den Testformaten mit offener oder freier Aufgabenbeantwortung diese selbst formulieren. Die Antworten können aus Einzelwörtern, mehreren Stichwörtern, aus Kurzantworten oder auch essayartigen Antworten bestehen. Die Problematik solcher Formate liegt weniger in ihrer Konstruktion als in der erreichbaren Objektivität bei der Be-

wertung der Antworten beziehungsweise der vergleichsweise großen Subjektivität der Bewertung.

Ergänzungsaufgaben (Lückentext)

Hierbei sollen Aussparungen in einem Text durch freie Antworten selbstständig gefüllt werden. Die geforderten Lösungen in den Lücken sind Wörter oder kurze Aussagen, die nach eindeutigen Kriterien als „richtig“ oder „falsch“ gewertet werden können. Vorteile dieses Testformats liegen in der Möglichkeit einer schnellen Abfrage von Wissenstatbeständen und in der Einfachheit der Korrektur. Zudem wird das Erraten einer Antwort oder das Auswählen scheinbar naheliegender Antworten vermieden, ebenso aber auch die „Bestrafung“ von Formulierungsschwächen. Allerdings können komplexe Lösungen mit diesem Testformat nur sehr bedingt erfragt werden.

Lückentexte sind nur dann den offenen Formaten zuzurechnen, wenn mehrere Möglichkeiten der Beantwortung richtig sind, sonst gelten sie als geschlossenes Format. Dies gilt auch, wenn nur bereits vorgegebene Begriffe zu ergänzen sind.

Assoziationsaufgaben

Die Schüler sollen auf vorgegebene Reize hin Assoziationen äußern („Nenne zentrale Begriffe / Namen aus dem Kontext XY!“). Aufgaben dieser Art sind gut anwendbar, wo eindeutige Zuordnungs-paare gebildet werden können. Bei einer vorhandenen Bandbreite der Antwortmöglichkeiten kann die Korrektur allerdings schwierig werden.

Substitutionsaufgaben

Der Schüler hat bestimmte, in der Aufgabe enthaltene Elemente durch neue, selbst gebildete Elemente zu ersetzen. So müssen in einem „Fehlertext“ eine oder mehrere falsche Aussagen als solche erkannt und verbessert werden. Eine gängige Version dieses Formats besteht darin, dass dem Schüler die Anzahl der zu korrigierenden Fehler bekannt ist.

Freiantwort-Aufgaben

Bei diesem Aufgabentyp werden kurze, möglicherweise nur einen Begriff oder eine Zahlenangabe oder wenige Worte umfassende Antworten verlangt. Üblich sind dabei Fragen beziehungsweise Auf-

forderungen des Typs „Wann war ...?“, „Nenne ...!“ oder „Was bedeutet ...?“.

Kurzaufsatz-Aufgaben

Dieser Aufgabentyp eignet sich besonders für die Darstellung komplexer Sachverhalte (Ursache-Folge-Zusammenhänge oder Entwicklungsverläufe) und die Darstellung von Problemen bis hin zu Beurteilungen und Bewertungen. Dieses Testformat, mit dem ohne Zweifel der Anforderungsbereich III angesprochen wird, erfordert allerdings einen relativ hohen Korrekturaufwand. Zur Erhöhung der Objektivität und der Validität dieses Testformats ist eine möglichst genaue vorherige Erstellung eines Erwartungshorizonts geboten.

Schritte bei der Konstruktion informeller Tests

Bei der Konstruktion informeller Tests sind im Wesentlichen folgende Einzelschritte zu durchlaufen:

- Bestimmung der Funktion des Tests
- Entscheidung, welche Lernziele Gegenstand des Test sein sollen
- Wahl des adäquaten Aufgabentyps
- Formulierung der Aufgaben (einschließlich der Schwierigkeitsanalyse, je nach Aufgabentyp: Trennscharfeanalyse, Distraktorenanalyse)
- Festlegung ihrer Reihenfolge im Testablauf
- Ausarbeitung einer Musterlösung
- Erstellen des Auswertungsschlüssels (Punktezuweisung)
- Normierung: Anlegen einer Benotungsskala

Umfang und Gestaltung

Tests können immer nur einen Ausschnitt dessen zum Gegenstand haben, was zuvor im Unterricht thematisiert worden ist. Tests sollten deshalb entsprechend dem Grundsatz der proportionalen Abbildung konstruiert werden. Die Auswahl dessen, was durch einen Test überprüft wird, darf nicht zufällig oder willkürlich, sondern soll systematisch und begründet erfolgen. Anders formuliert: was im Unterricht wichtig und bedeutsam war, sollte auch in einem Test einen entsprechenden Stellenwert erhalten.

Des Weiteren sollten den Schülern Tests in solchen Formaten vorgelegt werden, die zuvor im Unterricht bereits eingeübt wurden oder mit denen die Schüler

aus dem vorherigen Unterricht vertraut sind. Ein Test sollte nach Umfang und Schwierigkeitsgrad so gestaltet werden, dass er potenziell von allen Schülern einer Lerngruppe gelöst werden kann. Je nach Klassenstufe und inhaltlichem Anforderungsprofil sollte ein zeitlicher Umfang von 20 bis 30 Minuten nicht überschritten werden.

Die Aufgaben sollten unmissverständlich formuliert sein, keine Fallen, aber auch keine impliziten Lösungshinweise enthalten. Oft ist ein zusätzlicher deutlicher Hinweis auf die Art der Aufgabenbeantwortung hilfreich. Die Aufgaben sollten einerseits nicht zu leicht und damit unterfordernd sein; Tests sollten andererseits nicht zu viele Einzelaufgaben oder zu viele verschiedene Aufgabenformate enthalten. Bewährt hat sich eine Mischung aus Aufgaben mit gebundener sowie mit freier Aufgabenbeantwortung. Dabei sollten die Aufgaben den Schülern nicht zur Auswahl gestellt werden.

Aufgabenformate, die Kurzantworten in Stichworten, wenigen Sätzen oder in Form von Kurzaufsätzen erfordern, sind in ihrer Gewichtung grundsätzlich höher einzustufen als Aufgaben mit gebundener Aufgabenbeantwortung, zumal wenn in ihnen eine zumindest kurz gefasste Darstellung komplexer Zusammenhänge oder eine Beurteilungen mit knapper argumentativer Absicherung verlangt wird.

Auch im Fach Geschichte kommen für Testaufgaben nicht-verbale Darstellungsformen in Betracht. So können Aufgaben konstruiert werden, in denen Struktur-, Funktions- oder Prozess-Skizzen, Diagramme oder Schaubilder zu ergänzen, auszufüllen oder zu versprachlichen sind oder die die Umsetzung von Statistiken in grafische Darstellungsformen verlangen.

Es empfiehlt sich, nach Möglichkeit innerhalb eines Tests die zu lösenden Teilaufgaben in ihrem Schwierigkeitsgrad zu steigern. Um den Schülern eine Orientierung hinsichtlich der Wertigkeit der Teilaufgaben zu geben, sollte auf jeden Fall die jeweils zu erreichende Punktzahl (beispielsweise am Rand) angegeben werden.

Testauswertung

Die Bepunktung eines Tests erfolgt nach einem vorher erstellten (und für den Schüler beim Ausfüllen des Tests bekannten) Raster beziehungsweise einem zuvor formulierten Erwartungshorizont. Stellt sich eine Aufgabe nachträglich als problematisch oder gar unbrauchbar heraus, ist das Auswertungs-raster zu modifizieren. Verstöße gegen die allgemeine sprachliche Richtigkeit sollten grundsätzlich markiert, nicht aber in die Benotung (etwa bei freien Aufgaben) einbezogen werden, solange die sprachliche Richtigkeit nicht Gegenstand des Tests ist. Fachsprachliche Fehler dagegen sollten in angemessener Form in die Beurteilung einfließen. Es empfiehlt sich, eine Benotung mit „ausreichend“ bei Erreichen von etwa 45 % der maximal möglichen Punktzahl vorzunehmen. Die übrigen Notenstufen sollten dem Punktevolumen entsprechend dann in gleich großen Intervallen zugeordnet werden. Die Bestnote sollte demnach auch nicht erst bei Erreichen der Höchstpunktzahl erteilt werden. Hier erscheint eine Grenze von etwa 95 % sinnvoll.

Testbeispiele

Auf der Folgeseite finden Sie vier Beispieltests zum Thema „Preußische Reformen“, die in der Sekundarstufe I (Klasse 8) ein-

gesetzt werden können. Es handelt sich dabei um einen Multiple-Choice-Test (Test 1), einen Lückentext (Test 2), eine Kurzaufsatz-Aufgabe (Test 3) und eine Kurzaufsatz-Aufgabe (Test 4).

Test 1 ließe sich etwa mit 2 Punkten pro Aufgabe bewerten. Test 2 ließe sich mit einem Punkt pro Lücke, also 15 Punkten bewerten. Bei Test 3 könnte die Punktzahl pro Reformer und pro Maßnahme gegeben werden (etwa 4 Punkte pro Nennung). Bei Test 4 müsste die Bewertung anhand eines Erwartungshorizontes erfolgen, der eng auf den in der Klasse erreichten Unterrichtsstoff abzustimmen wäre.

LITERATUR

- Beutel, S.-I./ Vollstädt, W. (Hrsg.): Leistung ermitteln und bewerten. Hamburg 1999
- Borries, B. von: Test. In: Bergmann, K. u.a. (Hrsg.): Handbuch der Geschichts-didaktik. Seelze-Velber 1997 (5. Aufl.), S. 481 ff.
- Greving, J./ Linser, H./ Paradies, L.: Diagnostizieren und Fördern. Berlin 2008
- Greving, J./ Paradies, L./ Wester, F.: Leistungsmessung und -bewertung. Berlin 2005
- Rütter, T.: Formen der Testaufgabe. Eine Einführung für didaktische Zwecke. München 1973
- Sacher, W.: Tests und Klausuren in der Schule. Wie mache ich das? In: Beutel, S.-I./ Vollstädt, W. (Hrsg.): Leistung ermitteln und bewerten. Hamburg 1999, S. 63–74
- Weinert, F. E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim und Basel 2001

Test 1 (Multiple-Choice-Test) Reformen in Preußen

Name: _____

Klasse: _____

Kreuze jeweils die richtige Antwort an!

1. Durch das Edikt zur Bauernbefreiung von 1807

- verlor der preußische Adel seinen gesamten Landbesitz.
- wurde die Erbuntertänigkeit der Bauern aufgehoben.
- wurden alle Schulden der Bauern getilgt.
- wurde der preußische Adel gänzlich abgeschafft.

2. Durch die Städteordnung von 1808

- wurden die Rechte des vom König ernannten Bürgermeisters deutlich erweitert.
- wurde die von der Polizei zu kontrollierende Ordnung auf Straßen und Plätzen neu geregelt.
- wurde der Zunftzwang aufgehoben und die Gewerbefreiheit für jeden Bürger eingeführt.
- wurde die Selbstverwaltung der Städte durch die gewählten Stadtverordneten gewährleistet.

3. Durch die Heeresreform wurde

- das alte bewährte Söldnerheer personell verstärkt und neu bewaffnet.
- infolge der Einführung der allgemeinen Wehrpflicht ein Volksheer geschaffen.
- die zum Reglement der preußischen Armee zählende Prügelstrafe bewusst beibehalten.
- die Bewaffnung und Uniformierung der preußischen Armee dem französischen Vorbild angepasst.

4. Durch die Bildungsreform

- erhielten Frauen uneingeschränkten Zugang zu den Bildungseinrichtungen des Landes.
- sollten die Anlagen und Fähigkeiten eines jeden zum Wohle des Staates besser entfaltet werden.
- wurde die allgemeine Schulpflicht für jeden eingeführt.
- wurde in Preußen das Schulgeld abgeschafft.

5. Die preußischen Reformen waren die Voraussetzung für

- Preußens Aufnahme in den Rheinbund.
- die Wiederherstellung des Ansehens von König Friedrich Wilhelm III.
- eine Betonung der Standesunterschiede.
- Preußens wirtschaftlichen Aufschwung und militärische Stärke.

Test 2 (Lückentext) Reformen in Preußen

Name: _____

Klasse: _____

Fülle die Lücken in dem Lückentext:

Am Anfang der Reformen stand das Edikt über die Bauernbefreiung vom Oktober 1807. Es hob die Erbuntertänigkeit der Bauern auf und befreite damit den größten Bevölkerungsteil Preußens von der Jahrhunderte langen Bevormundung durch die adeligen Guts-herren. Die Bauern wurden persönlich frei, aber sie mussten weiterhin auf den Gutshöfen des Adels arbeiten und Abgaben leisten. Erst Hardenberg vollendete mit dem Regulierungsedikt von 1811 die Bauernbefreiung: Gegen Abtretung eines Drittels ihres Landes an ihre ehemaligen Herren wurden die Bauern volle Eigentümer und von Abgaben befreit. Aber vor allem den kleinen Bauern reichte das verbleibende Land zum Leben oft nicht aus. Sie verkauften es, wurden Landarbeiter oder zogen in die Stadt. 1810/11 wurde in Preußen der seit dem Mittelalter bestehende Zunftzwang aufgehoben und die Gewerbefreiheit eingeführt. Von nun an konnte sich jeder Bürger, der gegen eine Gebühr einen Gewerbeschein erworben und Gewerbesteuern bezahlt hatte, als Handwerker selbstständig machen.

Die neue Städteordnung vom November 1808 verwirklichte den Gedanken der städtischen Selbstverwaltung.

Test 3 (Kurzantwort-Aufgabe) Reformen in Preußen

Name: _____

Klasse: _____

Aufgabe:

Nenne vier preußische Reformer sowie die von ihnen vertretenen Reformmaßnahmen.

Test 4 (Kurzaufsatz-Aufgabe) Reformen in Preußen

Name: _____

Klasse: _____

Aufgabe:

■ Erläutere, warum die preußischen Reformen als „Revolution von oben“ bezeichnet wurden.

■ Nimm zu dieser Aussage in knapper Form begründend Stellung.