IEA (International Association for the Evaluation of Educational Archievment): Third International Math and Science Study (TIMSS). Boston College, Chestnut Hill, MA, USA 1997

Jude, Nina/Klieme, Eckhard: Sprachliche Kompetenz aus Sicht der pädagogisch-psychologischen Diagnostik. In: Beck, Bärbel/Klieme, Eckhard (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie. Weinheim: Beltz 2007, 9–22

Klieme, Eckhard: Bildungsqualität und Standards. Anmerkungen zu einem umstrittenen Begriffspaar. In: Friedrich Jahresheft XXIII (2005), 6–7

Klieme, Eckhard/Avenarius, Hermann/Blum, Werner/Döbrich, Peter/Gruber, Hans/Prenzel, Manfred/Reiss, Kristina/Riquarts, Kurt/Rost, Jürgen/Tenorth, Heinz-Elmar/Vollmer, Helmut J.: Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Bonn: BMBF 2003 (www.bmbf.de)

Nussbaumer, Markus: Lernerorientierte Textanalyse – Eine Hilfe zum Textverfassen? In: Feilke, Helmuth/Portmann, Paul R. (Hrsg.): Schreiben im Umbruch. Schreibforschung und schulisches Schreiben. Stuttgart: Klett 1996, 96–112

Portmann-Tselikas, Paul R.: Was müssen Lehrende über Aufgaben wissen? In: Bausch, Karl-Richard/Burwitz-Melzer, Eva/Königs, Frank G./Krumm, Hans-Jürgen (Hrsg.): Aufgabenorientierung als Aufgabe. Tübingen: Narr 2006, 182–191

Schneuwly, Bernard: Der Nutzen psychologischer Schreibforschung für die Didaktik des Schreibens. In: Feilke, Helmuth/Portmann, Paul (Hrsg.): Schreiben im Umbruch. Schreibforschung und schulisches Schreiben. Stuttgart: Klett 1996, 29–39

Weinert, Franz E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim, Basel: Beltz 2001

Becker-Mrotzek, M. (2014). Schreibleistungen bewerten und beurteilen. In H. Feilke & T. Pohl (Hrsg.), *Schriftlicher Sprachgebrauch – Texte verfassen* (S. 501–513). Baltmannsweiler: Schneider Verlag Hohengehren.

MICHAEL BECKER-MROTZEK

F2 Schreibleistungen bewerten und beurteilen

1 Einleitung

Das Bewerten und Beurteilen von Schülertexten stellt in mehrfacher Hinsicht eine Herausforderung dar, was sich zum einen herleitet aus der Komplexität von Texten und zum anderen aus der schulischen Praxis, die auf diese Herausforderungen nur bedingt angemessen reagiert. Im folgenden Artikel werden nach einer historischen und begrifflichen Erläuterung die Gründe für die Schwierigkeit der Textbewertung analysiert; anschließend werden die didaktischen Funktionen sowie Möglichkeiten des Bewertens und Beurteilens behandelt. Textbewertung – und das ist die zentrale Botschaft – kann und muss mehr sein als das Testen von Schreibkompetenz, das in F1 behandelt wird.

2 Historisches und Begriffliches

Schulische Leistungen und ihre Beurteilungen unterliegen einem historischen Funktionswandel (Huber 2002). Das heute dominierende Leistungsprinzip stellt grundsätzlich eine historische Errungenschaft dar, weil so seit dem frühen 19. Jh. der Zugang zu den Bildungseinrichtungen vom Prinzip her an die eigenen Anstrengungen gebunden ist, und nicht an Privilegien, soziale Herkunft oder andere zufällige Voraussetzungen. Es ist Ausdruck von Chancengleichheit. Allerdings erfährt es von Beginn an eine zweifache Verengung, die bis heute den Schulalltag bestimmt. Die erste betrifft die Art und Weise der Leistungs*erbringung*, die seit dem Ende des 19. Jh. – vor allem in den höheren Klassenstufen – als „schriftliche und mündliche Prüfungen von verbalisierten Fachkenntnissen (Klausuren und Examina)" realisiert werden (Huber 2002, 12). Die zweite betrifft die Form der Leistungs*beurteilung*, die immer mehr auf Ziffernnoten eingeschränkt wurde und andere Formen wie Verbalbeurteilungen verdrängte. Denn Ziffernnoten sind leicht zu erteilen und vor allem einfach zu verrechnen. Hierzu trägt aber auch die zunehmende Verrechtlichung der Leistungsbeurteilung bei, die ihrerseits zusammenhängt mit der Selektionsfunktion von Noten, über die der formale Zugang zu nachfolgenden Bildungsinstitutionen geregelt wird.

In der Fachliteratur findet sich eine verwirrende Vielzahl an Begriffen, von denen im Folgenden die hier interessierenden definiert werden sollen (vgl. Müller-Michaels 1993, 348f.):

- **Bewerten** ist ein kognitiver Akt des Einschätzens, der sich in unserem Fall auf Schülertexte bezieht. Dem Bewerten liegt – bewusst oder unbewusst – ein Wertmaßstab zugrunde, der sich in Form von Kriterien beschreiben lässt. Die Qualität einer Schreibleistung wird als „gelungen" oder „nicht gelungen" bewertet. Jeder Bewertungsvorgang ist für Lehrer wie Schüler mit Verstehensbemühungen verbunden, d. h. ihm geht immer ein Verstehen im hermeneutischen Sinn voraus. Bedeutet doch eine Bewertung im eigentlichen Sinn, „das Gute achten, den Wert betonen!" (Groß 1996, 18). Die Bewertung ist zugleich logische Grundlage für Revisionen aller Art, aber eben auch für das Beurteilen und Benoten.
- **Beurteilen** ist eine verbal geäußerte Bewertung, die hier gegenüber dem Schüler/Schreiber geäußert wird. Der bewertete Schülertext wird mehr oder weniger explizit bzw. implizit auf Normen, Kriterien und Wertmaßstäbe bezogen. Für die Schreibdidaktik ist zunächst einmal das „fördernde Beurteilen" bedeutsam (Baurmann 1996, 154), weniger das „prüfende Bewerten" (= Benoten). Mit der Beurteilung sollte das Bemühen um ein verantwortungsvolles Abwägen einhergehen. Verbale Beurteilungsäußerungen sollten als konstruktive Kritik und mit Hinweisen auf zukünftiges Verhalten formuliert sein, z. B. in Form konkreter Überarbeitungsvorschläge.

  Ein so verstandenes verbales Beurteilen findet in der Schule seinen Niederschlag in Lernentwicklungsbeschreibungen und Lehrerkommentaren unter Schülertexten. Diese Begriffsdefinition deckt sich mit vielen u. a. in den Richtlinien der Länder genannten pädagogischen Zielen. Die hier vorgestellte Form ist weitgehend eine Fremdbeurteilung. Zunehmend sollte diese in eine Selbstbeurteilung seitens der Schüler/innen (oder mindestens in eine Beteiligung der Lernenden an der Beurteilung) überführt werden.
- **Benoten** ist die zusammenfassende Bewertung einer Leistung in einer Ziffernote. Ihr kommt in der Schule traditionell und auch aktuell ein besonderes Gewicht zu. Vor dem Hintergrund der Bildungsstandards und zentralen Lernstandserhebungen nimmt deren Bedeutung, aber auch die Notwendigkeit einer kritischen Diskussion erheblich zu.

3 Texte bewerten

Die Aufsatzbeurteilung gilt seit langem als besonders problematisch, wie schon die Studie von Schröter (1971) über „Die ungerechte Aufsatzzensur" zeigt. Eine Replikationsstudie von Birkel (2003) belegt, dass die Beurteilungsvarianz bei Schülertexten nach wie vor groß ist, wobei sich vor allem im mittleren Leistungsbereich große Streuungen zeigen. Enthält ein Text zudem noch viele Rechtschreibfehler, ist die Note oft weniger Ausdruck der Textqualität als vielmehr der orthographischen Korrektheit. Insofern besteht in der Fachdidaktik seit längerem Einigkeit, dass Aufsatznoten keine Messungen im sozialwissenschaftlichen Sinn, sondern allenfalls Schätzungen sind, die die Leistungen innerhalb einer Klasse (oder anderen Bezugsgruppe) in eine Rangfolge bringen. Ihr Aussagewert gilt daher allgemein als gering, weil sie weder die Textqualität noch die Schreibkompetenz valide, reliabel und objektiv erfassen. Mit anderen Worten:

- Notenurteile sind nur bedingt gültig (valide), weil sie nicht die Textqualität, sondern eher das äußere Erscheinungsbild oder die orthographische Korrektheit erfassen.
- Sie sind nicht zuverlässig (reliabel), weil sie bei Wiederholungen nicht zum selben Ergebnis führen.
- Schließlich sind sie auch nicht objektiv, weil sie in starkem Maße davon abhängig sind, wer benotet.

Dessen ungeachtet kann die Schreibdidaktik aus sachlichen Gründen nicht auf das Bewerten von Texten verzichten, weil Schreiben immer auch Überarbeiten bedeutet. Überarbeiten ist aber in einer sinnvollen Weise nur möglich, wenn man eine Vorstellung davon hat, wie der angestrebte Text aussehen soll. Wir brauchen also Normen oder Kriterien, nach denen wir einen Text in seinen verschiedenen Aspekten bewerten und beurteilen, und zwar unabhängig davon, ob es um eigene oder fremde Texte geht. Von der Textqualität zu unterscheiden ist die Schreibkompetenz, die nicht ohne weiteres aus der Textqualität hergeleitet werden kann.

Texte als Mittel der schriftlichen Kommunikation stehen gewissermaßen zwischen Schreiber und Leser und stellen für beide eine Herausforderung dar. „Texte sind nicht einfach gegeben, sondern dem Leser aufgegeben", wie es Nussbaumer (1996, 98) treffend formuliert. Ohne das aktive Bemühen des Lesers um ein Verstehen wird aus dem **Text auf dem Papier** kein **Text im Kopf**. Insofern kann man einen Text danach beurteilen, wie gut oder schlecht er den Leser bei der Re-Konstruktion seines Sinns anleitet. Der Text auf dem Papier ist jedoch auch nicht identisch mit dem Text im Kopf des Schreibers, weil es uns nicht immer gelingt, unsere Ideen so in einem Text aufzuschreiben, wie wir uns das selber vorstellen. Daraus folgt, dass aus *einem* unzulänglichen Text nicht umstandslos auf die Ideen des Schreibers oder gar seine Schreibkompetenz geschlossen werden kann. Texte – insbesondere von Lernern – brauchen kooperative Leser, die bis auf Weiteres unterstellen, dass der Schreiber etwas mitzuteilen hat.

Mit dem Züricher Textanalyseraster (Nussbaumer 1996), das in einem didaktischen Forschungskontext entstanden ist und auf das sich zwischenzeitlich zahlreiche fachdidaktische Arbeiten beziehen, steht ein Bündel von linguistischen Merkmalen bereit, das sich zur Bewertung und Beurteilung von (Lerner-)Texten sehr gut eignet. Es unterscheidet fünf Dimensionen der Bewertung:

(0) **Grundgrößen:** Hierzu zählen – ohne Wertung – Eigenheiten wie die Textlänge, Wortschatz, syntaktische Muster oder Kohäsionsstiftung (Roter Faden).

(A) **Sprachformale Richtigkeit** (einschließlich Orthographie, Syntax und Semantik).

(B1) **Funktionale Angemessenheit:** Ist der Text auf dem Papier eine in Idee, Aufbau und verwendeten sprachlichen Mitteln angemessene Anleitung für den Leser, damit dieser ein Textverständnis bilden kann?

(B2) **Ästhetische Angemessenheit:** Ist der Text nicht nur verständlich, sondern genügt er darüber hinaus Ansprüchen für ein gutes Textdesign?

(B3) **Inhaltliche Relevanz:** Hat der Text zu seinem Thema etwas inhaltlich Bedeutsames zu sagen?

In Form einer Mind-Map lassen sich die Kriterien wie in Abbildung 1 darstellen. Diese Kriterien kann man als Suchanleitung lesen, um die sprachlich-textuellen Qualitäten eines Textes zu ermitteln – und zwar die positiven wie die negativen. Es ist kein festes Schema, sondern offen für die Anpassung an die je besonderen Bedürfnisse.

Aus einem in diesem Sinne gelungenen Text darf man in der Regel auf eine gut entwickelte Schreibkompetenz schließen. Denn gute Texte kommen nur selten zufällig zustande, weil sie in der Regel eine ausgebaute Schreibkompetenz erfordern. Aus weniger oder schlecht gelungenen Texten kann man nicht so einfach auf die Schreibkompetenz schließen, weil es hierfür unterschiedliche Gründe geben kann. Zum einen ist es möglich, dass der Schreiber ausnahmsweise einen schlechten Text verfasst oder auch die Aufgabe falsch verstanden hat. Zum anderen, und das ist gravierender, messen wir die Textqualität an gesellschaftlichen Erwartungen, die eine entwickelte Schreibkompetenz voraussetzen. Genau diese sollen Lerner aber erst erwerben, insofern benötigen wir für die Bewertung und Beurteilung von Lernertexten entwicklungssensible Kriterien, die auf das jeweilige Entwicklungsniveau abgestimmt sind.

4 Funktionen des Bewertens, Beurteilens und Benotens

Bewerten, Beurteilen und Benoten dienen insgesamt der Förderung der Schreibkompetenz. Gegenstand sind dabei sowohl die Schreibprozesse als auch die resultierenden Produkte, i.e. die Schülertexte. Denn aus der Schreibforschung wissen wir, dass Lerner nicht nur Rückmeldungen über die Qualität ihres Produktes benötigen, sondern auch Hinweise darauf, wie sie ihren Schreibprozess gestalten müssen, um gute Texte zu verfassen. Um diese wichtige pädagogische und didaktische Funktion des Bewertens und Beurteilens zu erfüllen, benötigen wir Verfahren, die den Anforderungen nach Transparenz, Validität, Objek-

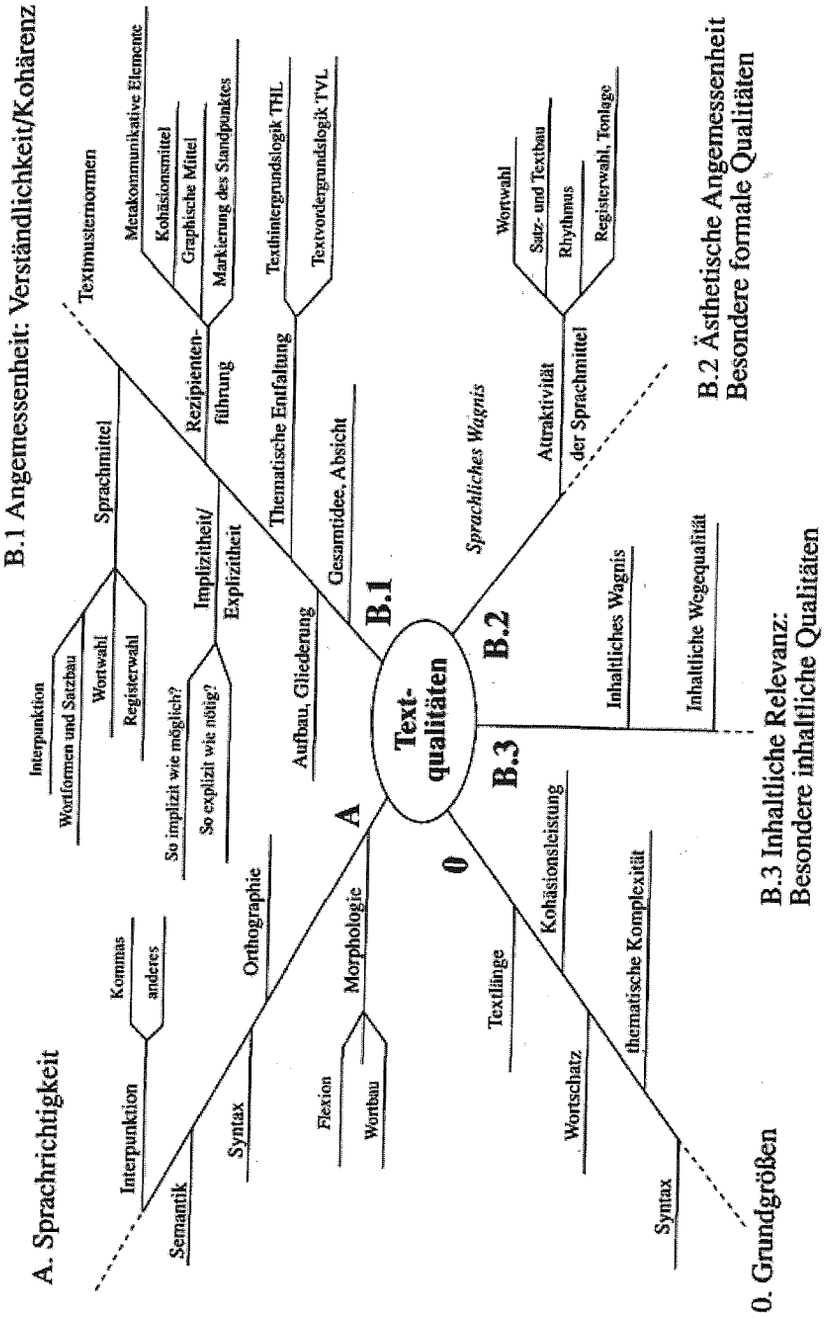

Abb. 1 Züricher Textanalyseraster (Nussbaumer 1996, 110)

tivität, Reliabilität und Fairness möglichst gerecht werden. Dabei kann und soll es nicht darum gehen, im Unterricht nur noch Aufgabentypen einzusetzen, wie wir sie aus den zentralen Lernstandserhebungen kennen (vgl. Becker-Mrotzek, F1). Denn das bedeutete ein Teaching-to-the-test, das niemand will. Es geht vielmehr um „eine schrittweise Verbesserung der Erfolgskontrolle des Schulunterrichts" (Groeben 2009, 116), indem die Praxis der Aufsatzbeurteilung systematisiert wird. Für Lehrende wie Lernende muss klar sein, nach welchen Kriterien Texte bewertet werden, und diese Kriterien sollen vielfältig sein.

Bei der Funktionsbestimmung des Bewertens ist zwischen Fremd- und Selbstbeurteilungen zu unterscheiden. Fremdbeurteilungen durch Lehrer, Schüler und Dritte stellen Rückmeldungen über den Gelingensgrad einer Schreibhandlung dar: Hat der Text sein Ziel erreicht? Rückmeldungen durch den Lehrer haben zwar den Vorteil, dass sie von einem Experten stammen, der auf die individuellen Lernbedingungen und -erfordernisse eingehen kann; zugleich sind sie aber immer auch eingebunden in die Notenfindung und damit die Selektionsfunktion. Rückmeldungen von Mitschülern sind in diesem Sinne weniger professionell-didaktisch, dafür aber authentisch und ohne Benotungsfunktion; das trägt wesentlich zu ihrer Akzeptanz bei. Aus diesem Grund ist es sinnvoll, im Unterricht beide Formen einzusetzen. Ein wesentliches Ziel von Fremdbeurteilungen ist die Entwicklung der Fähigkeit zur Selbstbeurteilung. Die Schüler und Schülerinnen lernen auf diese Weise, ihre Absicht und ihren Schreibplan mit ihrem Text kritisch zu vergleichen, um bei Abweichungen zu revidieren und zugleich künftige Schreibprozesse zu modifizieren. Mögliche Verfahren des Bewertens und Beurteilens werden weiter unten vorgestellt.

Das Benoten stellt wegen seiner Motivations- und Selektionsfunktion eine besondere Herausforderung dar und verlangt daher auch besondere Aufmerksamkeit in Bezug auf die Gütekriterien des Messens, insbesondere die der Transparenz und Fairness. Ebenso wie das Bewerten und Beurteilen beziehen sich Noten auf bestimmte Kriterien oder Normen. In der schulischen Praxis lassen sich insbesondere die folgenden drei Bezugsnormen unterscheiden:

- **Individuum als Bezugspunkt:** In diesem Fall werden die Leistungen der Lerner ins Verhältnis zu ihrem individuellen Entwicklungsstand gesetzt. Schreibleistung wird also bezogen auf zurückliegende Leistungen und damit in Relation zum Entwicklungsstand gesetzt; gemessen wird auf diese Weise der individuelle Kompetenzzuwachs (oder auch Kompetenzverlust).
- **Lerngruppe als Bezugspunkt:** Hierbei werden die Leistungen der Lernerinnen untereinander in Verhältnis gesetzt und damit in Relation zur mittleren Leistung der jeweiligen Lerngruppe gestellt. Auf diese Weise werden – testtheoretisch gesprochen – alle Lerner/innen in eine Rangfolge gebracht, so wie es auch bei standardisierten Tests der Fall ist. „Gerade diese Vergleichsperspektive hat aber dysfunktional destruktive Konsequenzen, die vom impliziten Konkurrenzdruck für die Schülerinnen und Schüler bis zu dem problematischen Zwang reichen, unabhängig vom Leistungsniveau einer Klasse bestimmte Lerner als leistungsschwach zu diagnostizieren" (Groeben 2009, 117). Genau dieser Bezugspunkt wird aber sehr häufig bei der Benotung von Klassenarbeiten zugrunde gelegt.
- **Inhaltliche Kriterien als Bezugspunkt:** Mit inhaltlichen Kriterien als Bezugspunkt haben wir es dann zu tun, wenn für die Beurteilung bestimmte sachbezogene Gesichtspunkte festgelegt werden. In diesem Fall werden die Leistungen eines Schülers bzw. einer Schülerin mit einem sachlichen Maßstab verglichen, so wie es etwa bei sportlichen Qualifikationen der Fall ist, wenn beispielsweise eine bestimmte Mindestzeit oder -weite erreicht werden müssen. Ein solches Verfahren ist die wohl angemessenste und konstruktivste Form der Beurteilung und Benotung, weil es prinzipiell allen Schüler/innen ermöglicht, das Lernziel zu erreichen. Zugleich handelt es sich um eine kompetenzorientierte Beurteilung und Benotung, so wie es die Bildungsstandards vorsehen. Voraussetzung hierfür sind jedoch explizite, d.h. möglichst operational definierte Lehr- und Lernziele, die dann in Form von Kriterienkatalogen vorliegen.

5 Verfahren des Beurteilens und Benotens

Im folgenden Abschnitt sollen möglichst konkrete Verfahren des Beurteilens und Benotens vorgestellt werden, mit deren Hilfe die oben genannten Funktionen realisiert werden können. Eine zentrale Rolle spielen dabei die sog. Kriterienkataloge, wie sie Becker-Mrotzek/Böttcher (2003) bzw. (2006) vorgestellt haben.

5.1 Kriterienkataloge

Kriterienkataloge stellen ein didaktisches Instrument dar, um die in einem Lernertext sichtbar werdenden Kompetenzen zu entdecken, zu bewerten und ggf. zu beurteilen. Hierfür benötigen wir zusätzliche, entwicklungssensitive Kriterien, die den jeweiligen Lern- und Entwicklungsstand berücksichtigen. Die Kriterienkataloge verfolgen das Ziel, die Anforderungen der Schreibaufgabe bzw. Textart mit den Entwicklungsmöglichkeiten der Lerner/innen zu verbinden. Für diesen Zweck werden die Dimensionen aus dem Züricher Textanalyseraster an das jeweilige Entwicklungsniveau angepasst, indem auf einem mittleren Abstraktionsniveau die jeweiligen Erwartungen formuliert werden. Dabei hat sich eine Anzahl von 10 bis 15 Kriterien als sinnvoll erwiesen, weil so jedem einzelnen Kriterium ein angemessenes Gewicht zukommt. Bei der Auswahl der Kriterien ist jedoch zu bedenken, dass nicht alle Kriterien die Anforderungen an eine angemessene Bewertung gleich gut erfüllen. Fragen nach der Gesamtstruktur sind beispielsweise relativ trennscharf, Frage nach der inhaltlichen Reichhaltigkeit dagegen nicht ohne Weiteres. In Anlehnung an das Züricher Textanalyse-

raster und das Modell von Baurmann (2002, 136) haben Becker-Mrotzek/Böttcher (2006, 92) einen Katalog mit zwölf Kriterien vorgestellt (vgl. die Abb. 2 weiter unten), die sich auf die Dimensionen „Sprachrichtigkeit", „Sprachangemessenheit", „Inhalt", „Aufbau" und „Schreibprozess" verteilen.

- Unter **sprachlicher Richtigkeit** werden insbesondere Orthographie einschließlich Zeichensetzung sowie Grammatikalität auf dem Stand der jeweiligen Klassenstufe bewertet.
- Unter **sprachliche Angemessenheit** fallen Umfang, Differenziertheit und gegebenenfalls Fachsprachlichkeit des Wortschatzes sowie der Satzbau, der in seiner Komplexität dem Thema und den Adressaten des Textes entsprechen soll.
- Unter **Inhalt** wird zum einen die Gesamtidee bewertet, d. h. die Frage beantwortet, ob die Idee zur Aufgabenstellung passt, sowie Umfang und Relevanz des geschriebenen Textes eingeschätzt.
- Unter **Aufbau** fallen insgesamt vier Kriterien: Hier gilt es zu beurteilen, ob das gewählte Textmuster angemessen ist, ob der Text sinnvoll, auch durch äußere Merkmale gegliedert ist, ob das Thema logisch und kohärent entfaltet wird und schließlich ob der Leser in irgendeiner Weise durch den Text geführt wird.
- Unter **Schreibprozess** fallen zum einen Planungs- und Überarbeitungsprozesse, die beispielsweise an Notizen und Überarbeitungsspuren zu erkennen sind; das bedeutet, dass entgegen einer weit verbreiteten schulischen Praxis Korrekturen im Text positiv bewertet werden – und nicht als mangelnde Ordnung. Und als besonderes Kriterium sind schließlich Wagnis und Kreativität angesetzt; hier besteht die Möglichkeit, gewagte, aber noch nicht erfolgreich realisierte Textelemente positiv zu bewerten.

Der Basiskatalog greift die für alle Textarten mehr oder weniger zutreffenden Kriterien auf. Er muss in zweifacher Weise an die jeweils konkreten Schreibaufgaben und Bewertungsziele angepasst werden: Zum einen müssen die spezifischen Merkmale der konkreten Aufgabe eingearbeitet werden, indem die einzelnen Kriterien konkreter gefasst werden. Und zum anderen gilt es, die didaktischen Ziele der Unterrichtseinheit zu berücksichtigen, indem beispielsweise einzelne Dimensionen durch zusätzliche Kriterien aufgewertet und andere durch Streichung reduziert werden.

Unter didaktischen und methodischen Gesichtspunkten ist es sicherlich sinnvoll, die Kriterienkataloge gemeinsam mit den Schülern zu erarbeiten, beispielsweise anhand von gelungenen und auch misslungenen Texten. Auf diese Weise werden sie für die Schüler/innen transparent und inhaltlich nachvollziehbar. In Leistungsüberprüfungen können sie in Form von Beurteilungsbögen den Schüler/innen schon mit der Aufgabenstellung ausgehändigt werden und von ihnen dann als Checklisten während der Arbeit benutzt werden.

Dimension	Kriterium	Grad		
Sprachrichtigkeit				
Orthographie	1. Entspricht die Orthographie einschließlich Zeichensetzung dem Lernstand?	1	0,5	0
Grammatikalität	2. Sind Wortbildung und Satzbau grammatisch korrekt?			
Sprachangemessenheit				
Wortwahl	3. Ist der Wortschatz angemessen? Werden Inhaltswörter, Funktionswörter, komplexe Ausdrücke und Fachtermini treffend verwendet?			
Satzbau	4. Ist der gewählte Satzbau der Aufgabe und dem Leser angemessen?			
Inhalt				
Gesamtidee	5. Lässt der Text eine dem Thema angemessene Gesamtidee erkennen (z. B. passende Überschrift)?			
Umfang/ Relevanz	6. Sind der Umfang und Inhalt der Aufgabe angemessen?			
Aufbau				
Textmuster	7. Wird ein der Aufgabe angemessenes Textmuster verwendet?			
Textaufbau	8. Ist der Text sinnvoll aufgebaut? Lässt er eine innere/äußere Gliederung erkennen (Abschnitte etc.)?			
Thematische Entfaltung	9. Wird das Thema in einer der Fragestellung angemessenen Art entfaltet?			
Leserführung	10. Wird der Leser aktiv durch den Text geführt? Werden Text strukturierende Mittel verwendet?			
Prozess				
Planen/Überarbeiten	11. Lässt der Text Planungs- und Überarbeitungsspuren erkennen?			
Wagnis/Kreativität	12. Lässt der Text ein besonderes sprachliches Wagnis erkennen? Ist er in besonderer Weise kreativ?			

Abb. 2 Basiskatalog zur Textbeurteilung

5.2 Benoten

Das Benoten von Texten stellt eine besondere Herausforderung dar, weil hier die „Notenfalle" besonders hart zuschlägt: „Die Ziffernnote kann der Komplexität eines geschriebenen Textes nicht gerecht werden" (Baurmann 2002, 120). Des-

halb muss eine prozess- und entwicklungsorientierte Schreibdidaktik zu einer neuen Beurteilungspraxis gelangen, die mindestens den folgenden Maximen verpflichtet ist:

- Förderung der individuellen schriftsprachlichen Fähigkeiten des einzelnen Schülers,
- Schülertexte als Entwürfe begreifen,
- Schreibprozess und Schreibprodukt gleichwertig nebeneinander stellen,
- das Beurteilen als dialogische, kommunikative Handlung betrachten,
- von der Fremdbeurteilung sukzessive zur Selbstbeurteilung gelangen.

Hierfür bieten sich Kriterienkataloge an, weil sie die Benotung transparenter, fairer und gerechter machen können. Durch die Notwendigkeit, für jedes einzelne Kriterium zu entscheiden, in welchem Grad es erreicht wird, werden Merkmale eines Textes sichtbar, die ansonsten unter den Mängeln in anderen Bereichen verborgen bleiben. Insbesondere orthographische Fehler und eine unleserliche Handschrift führen bei der Beurteilung häufig dazu, die inhaltliche Textqualität deutlich schlechter zu bewerten.

Wir haben für die Kriterienkataloge bewusst zwölf Kriterien gewählt, weil dies zur Folge hat, dass für die vier Notenstufen „sehr gut" bis „ausreichend" jeweils drei Punkte zur Verfügung stehen. Erst ein Text mit null Punkten würde danach als mangelhaft benotet. Bei der Bewertung von Klassenarbeiten sollte des Weiteren geprüft werden, ob es möglich ist, die Arbeit auf zwei aufeinanderfolgende Tage aufzuteilen. Am ersten Tag haben die Schüler Gelegenheit, den Text zu planen und zu formulieren; am zweiten Tag können sie mit einem gewissen Abstand ihren Text überarbeiten. Auf diese Weise wird der Schreibprozess systematisch in die Leistungsüberprüfung einbezogen. Eine weitere Möglichkeit besteht darin, die Texte nach dem ersten Tag zu kommentieren und den Schülern anschließend Gelegenheit zu geben, ihren Text zu überarbeiten.

5.3 Schreibaufgaben konzipieren

Eine weitere wichtige Voraussetzung für eine faire und transparente Beurteilung und Benotung stellen die Schreibaufgaben dar, die bestimmte Voraussetzungen erfüllen müssen, damit die Schüler/innen eine hinreichend klare Vorstellung über den erwarteten Text bilden können. Dazu muss die Aufgabe einerseits einen Handlungsrahmen aufspannen, in dem sinnvoll ein Text verfasst werden kann; andererseits muss die Aufgabenstellung dem Entwicklungsstand der Schüler und Schülerinnen entsprechen.

Ein Problem vieler Schreibaufgaben in Leistungsüberprüfungen besteht darin, dass für die Schülerinnen nicht erkennbar ist, welche Funktion der Text erfüllen soll. Das hängt damit zusammen, dass solche Aufgaben vielfach ohne erkennbaren Handlungszusammenhang und ohne auch nur vorstellbare Adressaten formuliert sind. In diesem Sinne handelt es sich um artifizielle, wenig authentische Schreibaufgaben, die praktisch keinen Bezug zu den alltäglichen Schreiberfahrungen der Schüler/innen haben. Wenn Schreiben jedoch eine problemlösende Handlung ist, wie es Hayes/Flower (1980) formulieren, dann ist es unverzichtbar, dass die Schreiber eine klare Zielvorstellung über den erwarteten Text ausbilden können. Denn es ist das Ziel, das jede Handlung steuert; ohne eine Zielvorstellung können wir keinen Handlungsplan bilden. Für die Konzeption von Schreibaufgaben bedeutet das, eine Situation zu entwerfen, aus der heraus sinnvoll ein Text geschrieben werden kann.

Wie eine solche schreibmotivierende Situation aussehen kann, lässt sich beispielhaft an einer Aufgabe der DESI-Studie zeigen (vgl. Beck/Klieme 2007). Eine der dort entwickelten Schreibaufgaben sieht vor, dass die Schüler/innen ein Reklamationsschreiben verfassen. Das Aufgabenblatt zeigt ein fiktives Sonderangebot für einen Computer, das unter anderem Ausstattungsmerkmale, Preis und Lieferbedingungen auflistet. Zusätzlich wird geschildert, dass beim Kauf des angebotenen Computers Abweichungen gegenüber dem Angebot aufgetreten sind, so wurden beispielsweise Lieferkosten berechnet. Die Schreibaufgabe besteht nun darin, ein förmliches Beschwerdeschreiben zu verfassen, in dem die Mängel benannt und Forderungen formuliert werden. Damit ist ein fiktives Szenario aufgespannt, das für fünfzehnjährige Schüler/innen einen nachvollziehbaren und vorstellbaren Schreibanlass darstellt. Zu den Merkmalen von guten Schreibaufgaben gehören mindestens die folgenden Merkmale (vgl. auch Baurmann/Pohl 2009, 100ff.):

- Ein **Handlungszusammenhang**, der einen Bezug zur Lebenswelt der Schüler/innen enthält und es den Schüler/innen ermöglicht, aus ihrem Wissen den Text inhaltlich zu füllen.
- **Informationen** über den Handlungszusammenhang, die einerseits mögliche Inhalte benennen und andererseits andere ausschließen.
- **Adressaten**, d. h. Leser, an die sich der Text richten soll.
- Eine **kommunikative Funktion**, die mit dem Text erreicht werden soll.

Eine besondere Schwierigkeit stellt das Einschätzen der Schwierigkeit einer Schreibaufgabe dar. Erste Hinweise für die Textart *Brief* liefert ebenfalls die DESI-Studie, die für drei Aufgabenmerkmale Ausprägungen benennt, die für die Schwierigkeit einer Schreibaufgabe verantwortlich sind (vgl. Neumann 2008). Bei den Merkmalen handelt es sich um den „erwarteten Text", den „Handlungskontext" sowie den „Impuls". Damit lassen sich für die Textart „Brief" die Anforderungen in die Schwierigkeitsgrade „einfach", „mittel" und „hoch" differenzieren:

- Eine Schreibaufgabe erweist sich als tendenziell **einfach**, wenn der Schreibimpuls, d. h. der Schreibauftrag, die erforderlichen Informationen möglichst einfach, einzeln und gegebenenfalls grafisch unterstützt präsentiert; wenn es sich bei dem Handlungskontext um einen individuellen, das heißt beispiels-

weise auf die Person des Schreibers bezogenen Zusammenhang handelt; und wenn der erwartete Text auch alltagssprachlich formuliert sein kann, beispielsweise in einer E-Mail oder SMS.

- Eine **mittelschwere** Schreibaufgabe zeichnet sich dadurch aus, dass der Schreibimpuls Einzelinformationen in Textform präsentiert; dass es sich beim Kontext um einen gruppenbezogenen Handlungsrahmen handelt; und dass der erwartete Text die Beachtung textartspezifischer Normen voraussetzt.
- Bei einer **schweren** Schreibaufgabe wird der Schreibimpuls in komplexer Textform präsentiert; der Kontext besteht aus einem größeren, gesellschaftlichen Handlungsrahmen; und der erwartete Text verlangt eine eigenständige Problemlösung sowie die Beachtung allgemeiner Standards, wie sie beispielsweise in der DIN-Norm für Briefe niedergelegt sind.

Diese Merkmale sollen an einem Beispiel kurz illustriert werden: Wenn das Schreibszenario die Schließung einer Sporteinrichtung vorsieht und der Schreibauftrag lautet, hierzu eine Stellungnahme zu verfassen, dann ist das einfacher, wenn der Impuls neben einem kurzen Text beispielsweise auch ein Bild mit einem geschlossenen Schwimmbad und davor wartenden Jugendlichen zeigt und der Schreibauftrag darin besteht, einen Freund in einer E-Mail darüber zu informieren. Deutlich schwieriger ist es, wenn der Impuls aus einem längeren Text besteht, in dem die Gründe für die Schließung dargelegt werden, und der Schreibauftrag vorsieht, stellvertretend für die Jugendlichen einen Brief an die Stadtverwaltung zu schreiben.

5.4 Weitere Bewertungs- und Beurteilungsverfahren

Neben die gewohnten Verfahren der Leistungsüberprüfung sollten weitere treten, um den Schülern bei unterschiedlichen Schreibanforderungen Gelegenheit zu bieten, ihre Fähigkeiten zu zeigen. Denn die traditionelle Klassenarbeit ist eine sehr spezifische, wenig authentische und auf einige Anforderungsbereiche eingeschränkte Form der Leistungsüberprüfung. Als alternative Form ist insbesondere das Portfolio zu nennen. Portfolios sind Mappen, in denen Schüler/innen nach vorher festgelegten Kriterien ihre schriftlichen Produkte und Reflexionen sowie weitere Materialien sammeln. Sie geben damit zugleich Auskunft über den Entwicklungsprozess, und zwar sowohl für die Lerner als auch für die Lehrenden. Dabei gilt, dass es unterschiedliche Möglichkeiten für das Führen und Bewerten von Portfolios gibt, die mit Blick auf die Lerngruppe, die didaktischen Ziele sowie die verfügbare Zeit zu entwickeln sind.

Exemplarisch sei hier auf eine Idee von Bräuer (1998) verwiesen, der zwischen Produkt- und Prozessportfolios unterscheidet. Insbesondere die Prozessportfolios bieten sich für einen entwicklungsorientierten Schreibunterricht an. In ihnen sammeln die Schüler und Schülerinnen über einen bestimmten Zeitraum ihre Schreibprodukte, und zwar von den ersten Entwürfen bis zum fertigen Ergebnis,

so dass der Entwicklungsprozess sichtbar wird. Am Ende entscheiden sie, welche Texte für die Beurteilung und Benotung herangezogen werden sollen (vgl. auch Becker-Mrotzek/Böttcher 2006, 104 ff.).

Literatur

Baurmann, Jürgen: Geschriebenes beurteilen. In: Praxis Deutsch, Sonderheft Schreiben (1996), 154–155

Baurmann, Jürgen: Schreiben, Überarbeiten, Beurteilen. Ein Arbeitsbuch zur Schreibdidaktik. Seelze: Kallmeyer 2002

Baurmann, Jürgen/Pohl, Thorsten: Schreiben – Texte verfassen. In: Bremerich-Vos, Albert/Granzer, Dietlinde/Behrens, Ulrike/Köller, Olaf (Hrsg.): Bildungsstandards für die Grundschule: Deutsch konkret. Berlin: Cornelsen 2009, 75–103

Beck, Bärbel/Klieme, Eckhard (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messungen. DESI-Ergebnisse. Band 1. Weinheim: Beltz 2007

Becker-Mrotzek, Michael/Böttcher, Ingrid: Schreibkompetenz entwickeln und beurteilen. Berlin: Cornelsen 2006

Birkel, Peter: Aufsatzbeurteilung – Ein altes Problem. In: Didaktik Deutsch (2003), Heft 15, 46–63

Böttcher, Ingrid/Becker-Mrotzek, Michael: Texte bearbeiten, bewerten und benoten. Unter Mitarbeit von Carolin Speckgens. Berlin: Cornelsen Scriptor 2003

Bräuer, Gerd: Portfolios. Lernen durch Reflektieren. In: Informationen zur Deutschdidaktik 22 (1998), Heft 4, 80–91

Groeben, Norbert: Möglichkeiten und Grenzen des Testens kultureller Bildung. In: Bertschi-Kaufmann, Andrea/Rosebrock, Cornelia (Hrsg.): Bildungsaufgabe und Forschungsfeld. Weinheim, München: Juventa 2009, 107–124

Groß, Barbara: Das Gute achten. Bewertung freigestalteter Arbeitsprodukte. In: Friedrich-Jahresheft XIV (1996), 18–19

Hayes, John R./Flower, Linda: Identifying the organisation of writing processes. In: Gregg, Lee W./Steinberg, Erwin (Hrsg.): Cognitive processes in writing. Hillsdale, N. J.: Erlbaum 1980, 3–30

Huber, Ludwig: Leistungen in der Schule. Rückblicke in die Geschichte – Fragen an die Gegenwart. In: Winter, Felix/Groeben, Annemarie von der/Lenzen, Klaus-Dieter (Hrsg.): Leistung sehen, fördern, werten. Neue Wege für die Schule. Bad Heilbrunn: Klinkhardt 2002, 11–19

Müller-Michaels, Harro: Noten für Kreativität? Zum Problem der Beurteilung produktiver Arbeiten im Literaturunterricht. In: Deutschunterricht (Berlin) 46 (1993), Heft 7/8, 338–348

Neumann, Astrid: Textproduktion in der Sekundarstufe I. Empirische Hinweise zur vergleichbaren Messbarkeit von Schreibleistungen aus der Sicht des Large-Scale-Assessments DESI. In: Bremerich-Vos, Albert/Granzer, Dietlinde/Köller, Olaf (Hrsg.): Lernstandsbestimmungen im Fach Deutsch. Gute Aufgaben für den Unterricht. Weinheim, Basel: Beltz 2008, 117–132

Nussbaumer, Markus: Lernerorientierte Textanalyse – Eine Hilfe zum Textverfassen? In: Feilke, Helmuth/Portmann, Paul R. (Hrsg.): Schreiben im Umbruch. Schreibforschung und schulisches Schreiben. Stuttgart: Klett 1996, 96–112

Schröter, Gottfried: Die ungerechte Aufsatzzensur. Bochum: Kamp 1971