

Bibliographischer Hinweis sowie Verlagsrechte bei den online-Versionen der DD-Beiträge:



**Halbjahresschrift für die Didaktik
der deutschen Sprache und
Literatur**

<http://www.didaktik-deutsch.de>
13. Jahrgang 2008 – ISSN 1431-4355
Schneider Verlag Hohengehren
GmbH

Juliane Köster & Thomas Lindauer

**ZUM STAND
WISSENSCHAFTLICHER
AUFGABENREFLEXION AUS
DEUTSCHDIDAKTISCHER
PERSPEKTIVE**

In: Didaktik Deutsch. Jg. 13. SH. 2008.
S. 148-161.

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten. Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder andere Verfahren – reproduziert oder in eine von Maschinen, insbesondere von Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. – Fotokopien für den persönlichen und sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopien hergestellt werden.

Juliane Köster, Thomas Lindauer

ZUM STAND WISSENSCHAFTLICHER AUFGABENREFLEXION AUS DEUTSCHDIDAKTISCHER PERSPEKTIVE

Die KMK-Bildungsstandards für das Schulfach Deutsch haben nicht nur eine Debatte über Standards ausgelöst, sondern die Deutschdidaktik auch mit zwei eng verbundenen Diskussionsfeldern konfrontiert: dem Feld der Kompetenzmodelle und dem der Aufgaben.

Folgt man dem Diskurs wissenschaftlicher Bemühungen, dann werden Aufgaben zum einen entwickelt, um die Bildungsstandards zu konkretisieren. Zum anderen sollen auf der Basis der Aufgabenlösungen lernbereichsspezifische, d. h. für Sprachrezeption (Hören, Lesen), Sprachproduktion (Sprechen, Schreiben) sowie für Sprachreflexion jeweils unterschiedliche Kompetenzmodelle etabliert bzw. modifiziert werden.¹ Allerdings erscheinen die Probleme mit kohärenten und stringenten Kompetenzstufenmodellen gegenwärtig noch so erheblich, dass zunächst einfachere Module von Kompetenzen eine realistische Zielgröße darstellen könnten.

1. Wer trägt zum Aufgabendiskurs bei und wer bestimmt ihn?

Wie im gesamten fachdidaktischen Feld so gibt es auch im Bereich „Aufgaben“ vier Einflussgrößen: die unterrichtliche Praxis, die Bildungswissenschaften im weitesten Sinn – einschließlich wissenschaftlicher Fachdidaktiken –, die Bildungsadministration und die Bildungsmedien.

1.1. Die unterrichtliche Praxis

Die unterrichtliche Praxis stellt selbst bereits ein Integrationsfeld dar, in dem in durchaus unterschiedlicher Weise „Brauchtum“, Erträge fachdidaktischer Forschung, bildungsadministrative Setzungen und Angebote der Bildungsmedien – mehr oder weniger reflektiert – wirksam werden. Die unterrichtliche Praxis ist im vorliegenden Zusammenhang nicht allein deshalb von großer Bedeutung, weil sie generell Zielgröße fachdidaktischer Reflexion ist, sondern weil sie bezüglich Planung und Durchführung in erheblichem Maße durch die „Aufgabe“ als „*movens agens*“ bestimmt ist (Lechner 2007, S. 127). So gilt auch für die Ausbildung in den Studienseminaren, dass „[d]er Unterricht [...] *nicht* durch Präambeln und Standards verbessert [wird], sondern durch ‚gute‘ Aufgaben, die im Klassenzimmer ankommen“ (<http://www.aufgabenkultur.studienseminar-koblenz.de>).

Während die unterrichtliche Praxis sich gegenüber den Einflüssen von Fachdidaktik und Bildungsmedien in unterschiedlichem Maß öffnen kann, sind bildungsadministrative Maßnahmen für sie verpflichtend und kommen im Bereich der „Aufgabenkultur“ traditionsgemäß stark zur Geltung. Für eine wissenschaftliche Aufgabendi-

¹ Vgl. Ossner 2006 sowie die Diskussion dieses Modells in *Didaktik Deutsch* 22 (2006), S. 5-24.

daktik ist die unterrichtliche Praxis vor allem deshalb interessant und wichtig, weil Lehrpersonen über reiche Erfahrung verfügen und wissen, welche Aufgaben auf welche Weise funktionieren. Dieses *Wissen* ist für die Deutschdidaktik auch dann bedeutsam, wenn Fachdidaktiker die Beobachtungen der Fachlehrer anders interpretieren als diese selbst. Worauf es ankommt, ist die Wertschätzung der Lehrer- und Lernerperspektive in einem Sektor, auf dem sich vor allem diese beiden Gruppen abarbeiten.

1.2. Die Bildungsadministration

Dass in der unterrichtlichen Praxis im Aufgabenbereich seit Langem zwischen „Wiedergabe“, „Anwendung“ und „Transfer“ unterschieden wird, verdankt sich bildungsadministrativen Setzungen. Diese deutlich zu unterscheidenden kognitiven Aktivitäten prägen seit den 70er-Jahren die einheitlichen Bestimmungen für die Abiturprüfung (http://www.kmk.org/doc/beschl/epa_deutsch.pdf, S. 11 ff.) in allen Fächern und sind als Anforderungsbereiche (I bis III) auch in die nationalen Bildungsstandards im Fach Deutsch (www.kmk.org/schul/Bildungsstandards/Deutsch_MSA_BS_04-12-03.pdf) eingegangen.

Der Einfluss der Bildungsadministration auf die Aufgabenkultur wächst in Zeiten Zentraler Abschlussprüfungen und Assessments zur Bilanzierung von Lernergebnissen. Je sichtbarer die administrativen Vorgaben – z. B. für zentrale Prüfungen und Schulleistungstests – in der schulischen Praxis sind, desto zugänglicher werden sie für die wissenschaftliche Analyse und Reflexion. Entsprechendes gilt natürlich auch für die schulische Rezeption der administrativen Verordnungen.

Analysiert man die administrativen Setzungen, dann wird zum einen deutlich, wie stark sie z. B. in Bezug auf die drei Anforderungsbereiche und Operatoren mit schulischem Brauchtum verquickt sind – z. B. wenn es um Offenheit und Komplexität der Aufgaben geht oder darum, dass Textverstehen in Form schriftlicher Produktionsleistungen zu überprüfen sei.²

Zum anderen orientiert sich die Bildungsadministration auch an den von der Psychologie, der Psychometrie, der Informatik und der Testtheorie bevorzugten Aufgabenkonzepten. Denn die KMK-Amtschefskommission „Qualitätssicherung in Schulen“ hat in Verbindung mit der Arbeit an den Bildungsstandards die Gründung eines Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) beschlossen, das die Länder großzügig finanzieren. Diese Tatsache zeigt deutlich, dass die Bildungsadministration keineswegs nur auf die traditionellen Überprüfungsformen setzt, sondern ebenso stark auf „die Entwicklung, Auswertung und kompetenzdiagnostische Rückmeldung von standardbasierten Testaufgaben“ (<http://www.iqb.hu-berlin.de/>), die ganz anderen Gesetzen folgen als die Aufgabenbeispiele in den Bildungsstandards und den EPA Deutsch.

² Vgl. z. B. folgende Passage aus den Bildungsstandards Deutsch für den Mittleren Schulabschluss: „Die Bearbeitung der Aufgaben erfordert die Fähigkeit, Schreibprozesse wie [sic!] mündliche Aussagen zu gestalten, zu kommunizieren, mit Texten umzugehen, ihnen Informationen zu entnehmen und Sachverhalte und Problemstellungen angemessen zu artikulieren.“ (S. 19) Vgl. auch die konkreten Aufgabenbeispiele.

1.3. Die Bildungswissenschaften

Zu den mit „Aufgaben“ befassten Wissenschaften zählen neben den psychologischen Disziplinen (Lern- und Kognitionspsychologie, Psychometrie) in erster Linie die Fachdidaktiken unter Einschluss ihrer Bezugswissenschaften.

Von psychologischer Seite sind vor allem die Kognitionspsychologie und die Psychometrie in durchaus unterschiedlicher Funktion in dieses Arbeitsfeld involviert. Am Beispiel einiger PISA-Aufgaben zeigen Schnotz/Dutke (2004, S. 61f.), dass Kognitionspsychologie und Psychometrie durchaus unterschiedliche Interessen verfolgen. Während Kognitionspsychologen die geistigen Aktivitäten beim Lösen einer Aufgabe identifizieren, geht es den Psychometrikern v. a. darum, Leistung – also z. B. auch geistige Aktivitäten – messbar zu machen.

Unter den Fachdidaktiken engagieren sich sowohl die mathematisch-naturwissenschaftlichen als auch die fremdsprachlichen Disziplinen im Bereich der Aufgabenreflexion. Die Mathematikdidaktik³ befasst sich – angestoßen durch die TIMSS-Resultate – seit den 90er-Jahren intensiv und differenziert mit Fragen der Aufgabenbeschaffenheit. Auch die Fremdsprachendidaktiken⁴, deren Vorlauf sich dem Gemeinsamen Europäischen Referenzrahmen (GER) verdankt, unterhält eine ausgeprägte Aufgabendebatte, die durch internationale Impulse geprägt ist. Das gilt auch für Deutsch als Fremd- bzw. Zweitsprache (DaF/DaZ), eine Disziplin, die der Aufgabenkonstruktion traditionell viel Aufmerksamkeit widmet und für die Deutschdidaktik vielfältiges Anregungspotenzial bietet (vgl. Häussermann/Piepho 1996; Funk 2006, S. 52-61).

Gemessen an der gesellschaftlichen Bedeutung der deutschdidaktischen Disziplin ist deren Beitrag zur Aufgabenanalyse und -reflexion noch zurückhaltend. Erst seit 2003 liegen theoretische und empirisch-qualitative Analysen vor, die auf konkrete Aufgabenmerkmale und auf Aufgabennutzung gerichtet sind. Die Arbeiten beziehen sich auf die Unterscheidung von Lern- und Leistungsaufgaben (vgl. Köster 2003; Lindauer/Schneider 2007), auf den Schwierigkeitsgrad von Leseaufgaben (vgl. Schweitzer 2007) und auf die Nutzung von Aufgabenangeboten in Lehrwerken (vgl. Winkler 2005, S. 177-196). Gegenwärtig ist zudem zu registrieren, dass die Bildungsadministration nicht nur bei der Entwicklung von Testaufgaben deutschdidaktische Beratung sucht (wie z. B. am IQB oder bei VERA), sondern auch bei der Entwicklung und Evaluation von Aufgaben für Zentrale Prüfungen (so z. B. NRW). Um solche Beratungsaufträge verantwortlich, d. h. wissenschaftlich einigermaßen gesichert wahrnehmen zu können, braucht es innerdisziplinäres gemeinsames Forschungsbemühen und kritische Debatte.

³ Vgl. die Arbeiten von Werner Blum, Regina Bruder, Timo Leuders und Michael Neubrand.

⁴ Z. B. das Forschungs- und Nachwuchskolleg „Lernaufgabenforschung in schulischen Kontexten“. Eine Kooperation zwischen den Pädagogischen Hochschulen Freiburg und Heidelberg. Vgl. auch Legutke 2006, S. 140-148.

1.4. Die Bildungsmedien

In dem Maß, in dem sich Deutschdidaktik in der Aufgabenforschung zurückhält, springen die Bildungsmedien in die Bresche und machen – im Rückgriff auf bildungsadministrative Setzungen – mehr oder weniger reflektierte Angebote. Sie tun das durchaus auch unter Beteiligung fachdidaktischer Expertise. Die Bildungsmedien können im Idealfall erfahrene und innovative Lehrkräfte für die Aufgabenentwicklung verpflichten und sie – unter der Ägide engagierter Redakteurinnen – mit Didaktikern ins Gespräch bringen. Das Resultat sind Leitmedien, die nicht nur die unterrichtliche Praxis maßgeblich (und mitunter durchaus inspirierend) bestimmen, sondern auch bildungsadministrative Entscheidungen in nicht unerheblichem Maß beeinflussen.

Fazit: Es stellt sich also insgesamt die Frage, ob die Deutschdidaktik als wissenschaftliche Disziplin maßgeblich am Aufgabendiskurs beteiligt sein will. Denn die aufgabenbezogenen Angebote aus dem Bereich der Bildungsmedien und der Bildungsadministration sind prominente Objekte deutschdidaktischer Analyse und Reflexion. Hier findet Deutschdidaktik neben der Psychometrie ein Feld, auf dem sie nicht kompetent ersetzt werden kann. Aus analytisch-evaluierender Perspektive hätte sie – bei aller Kooperation – auch Leistung und Grenzen testmethodischer Arrangements zu diskutieren. Das gilt auch – im Hinblick auf die traditionellen Prüfungsaufgaben – für das Verhältnis zur Bildungsadministration. Denn ohne konstruktiv-kritische Zusammenarbeit mit den für die Testmethodik verantwortlichen Wissenschaftlern einerseits und den die Tradition repräsentierenden Experten in den Kultusministerien und Qualitätsagenturen der Länder andererseits erscheinen Lösungen der schwierigen Fragen, die sich im Spannungsfeld von Multiple-Choice-Aufgaben und offenen Erschließungsformen stellen, kaum zu realisieren.

2. Zur Forschungslage

Vor diesem Hintergrund steht die deutschdidaktische Aufgabenforschung zunächst vor zwei Fragen: Worauf kann sie zurückgreifen und woran kann sie anschließen? Die folgende Bestandsaufnahme rekurriert auf Ergebnisse der oben genannten Wissenschaften und bestätigt damit, dass Aufgabenreflexion ein interdisziplinäres Unternehmen ist.

2.1. Grundlegende definitorische Unterscheidungen

- Unterscheidung von Aufgabenanforderungen und Lerneraktivitäten

Aufgaben integrieren Leistungsanforderungen und Lerneraktivitäten. Diese Unterscheidung resultiert aus dem Jenkins'schen Modell (Jenkins 1979), das die für Lernprozesse relevanten Komponenten abbildet und das – wie Artelt (Artelt u. a. 2005, S. 12) für den Bereich der Lesekompetenz gezeigt hat – auch Aufgaben in anderen Subdomänen determiniert.

Während die *Leistungsanforderungen* primär durch den Lern- bzw. Prüfungsgegenstand bestimmt sind, handelt es sich bei den *Aktivitäten der Lerner* v. a. um die von diesen zu erbringende kognitive – und ggf. auch den materielle – Aktivität (z. B.

identifizieren, strukturieren, explorieren). Zu den aufgabenrelevanten *Lernermerkmalen* gehören die für die Aufgabenbewältigung vorausgesetzten Wissensbestände (textbezogenes Wissen, kulturelles Wissen, rhetorisch-stilistisches Wissen, grammatisches, lexikalisches Wissen etc.) und Skills. Gegenstandsseitig spielt dessen *Beschaffenheit*, d. h. die Komplexität der Struktur des zu rezipierenden oder zu produzierenden *Gegenstands*, eine den Schwierigkeitsgrad der Aufgabe beeinflussende Rolle.

- Unterscheidung von Lern- und Leistungsaufgaben

Der *Erwerb* von Wissen und Können folgt anderen Gesetzen als deren *Überprüfung*. Diese Unterscheidung von Aufgaben ist populär, seit Weinert (1999) als Ergebnis von ihm betreuter Längsschnittprojekte auf die „völlig unterschiedlichen psychologischen Gesetzmäßigkeiten“ von „Lernen und Leisten“ verwiesen hat, auch wenn „Leistungen in der Regel vom vorhergehenden Lernen abhängen“. Eine Aufgabe kann für die Lernsituation überaus geeignet sein und für die Leistungssituation denkbar ungünstig. Denn in einer Leistungssituation bemüht man sich in erster Linie darum, Erfolge zu erzielen und Misserfolge zu vermeiden. Demgegenüber geht es in Lernsituationen darum, Neues zu lernen, Wissenslücken zu schließen oder unklar Gebliebenes besser zu verstehen. Während es in der Leistungssituation Fehler zu vermeiden gilt, sind sie in der Lernsituation ein Erkenntnismittel. Erfolgreicher Unterricht braucht nach Weinert beides, und in der Wahrnehmung der Schüler möglichst getrennt: viele entspannte Gelegenheiten zum intensiven Lernen und genügend anspruchsvolle Leistungssituationen. Folglich erscheint es sinnvoll, für beide Situationen unterschiedliche Aufgaben zu modellieren. (Für eine eingehende Diskussion zur Unterscheidung und Differenzierung von Lern- und Leistungs- bzw. Prüfungsaufgaben vgl. Köster 2003; Legutke 2006; Lindauer/Schneider 2007; Schneider/Lindauer 2007.)

2.2. *Beschreibungskategorien – Aufgabenmerkmale*

- Aufgabenformate und Aufgabenarten

Spätestens seit dem Bericht über PISA 2000 ist die prinzipielle Unterscheidung von geschlossenen, halboffenen und offenen Aufgaben auch in der deutschdidaktischen Diskussion präsent. Zu Beginn der 70er-Jahre wurde diese Debatte v. a. in der Auseinandersetzung mit dem „Aufsatz“ als Prüfungsform von Ingenkamp (1971) angestoßen und stellt für die Deutschdidaktik vor allem deshalb eine Herausforderung dar, weil die Offenheit einer Aufgabe per se als Qualitätsmerkmal gilt.

2003 hat Bremerich-Vos in der Zeitschrift „Deutschunterricht“ die diesen Formaten zugeordneten Aufgabenarten vorgestellt (vgl. Bremerich-Vos 2003). Zu den geschlossenen Aufgabenformaten zählen: Richtig-/Falsch-Items, Multiple-Choice-Items, Zuordnungs-Items, Umordnungs-Items.

Als halboffen gelten Kurzantwort-Items und – bei Schreibaufgaben – Lückentexte. Überall da, wo die Lerner kleinere oder größere Texte mündlich oder schriftlich produzieren, handelt es sich um offene Formate.

- Aufgabenschwierigkeit

Während der empirische Schwierigkeitsgrad einer Aufgabe aus der Anzahl richtiger Lösungen resultiert (je geringer die Häufigkeit richtiger oder plausibler Lösungen, desto höher der empirische Schwierigkeitsgrad), muss der theoretische Schwierigkeitsgrad kategoriengeleitet bestimmt werden. Auch hier hat die Psychometrie im Anschluss an PISA 2000 hilfreiche Margen gesetzt. Artelt (Artelt/Stanat/Schneider/Schiefele/Lehmann 2004, S. 154-157) unterscheidet auf der Basis der PISA-Leseaufgaben drei schwierigkeitsbestimmende Merkmale: den *Entscheidungsspielraum*, den eine Aufgabe dem Lerner bietet bzw. zumutet, den *Integrationsgrad*, den die Aufgabenbewältigung impliziert, und den *Präzisionsgrad* als Maß für die Genauigkeit, die mit der Aufgabenbewältigung verbunden ist – sei es auf methodisch-strategischer, sei es auf inhaltlicher Ebene. Dass diese drei differenzierenden Aufgabenmerkmale deutschdidaktischer Reflexion dienlich sein können, zeigt deren Rezeption in deutschdidaktischen Forschungsprojekten (Schweitzer 2006). Das ist v. a. auch darin begründet, dass die Merkmale „Entscheidungsspielraum“ und „Integrationsgrad“ große Anbindungspotenziale für die im didaktischen Zusammenhang hochwertigen Kategorien „Offenheit“ und „Komplexität“ bieten. Beide Merkmale bestimmen das kognitive Anregungspotenzial einer Aufgabe und verlangen deshalb auch fachdidaktische Reflexion. Geht es doch darum, optimale Stufungen und Verhältnisbestimmungen von Offenheit und Komplexität zu gewinnen (vgl. Köster 2005), wobei aus deutschdidaktischer Perspektive auch Aspekte der konkreten Aufgabenformulierung (vgl. Schäfers 2006) eine Rolle spielen. Will man den Forschungsstand unter Einschluss deutschdidaktischer Arbeiten resümieren, dann geht es vor allem darum, welche kognitiven Aktivitäten eine Aufgabe erfolgreich anstößt, sodass die mit der Aufgabe verbundene Anforderung eine dem Gegenstand und dem Lerner angemessene Bewältigung erfährt. Zu große Offenheit – ggf. im Verbund mit zu großer Komplexität – führt bei schwachen Lernern zu Verunsicherung und Ratlosigkeit, bei mittlerer Leistungsfähigkeit zu „wilder Komplexitätsreduktion“ durch Beschränkung auf Bekanntes oder den Rückgriff auf vertraute Routinen (Köster 2006a). Der Präzisionsgrad von Aufgaben als drittes schwierigkeitsbestimmendes Merkmal mag demgegenüber speziell für Testaufgaben und hier wiederum für die PISA-Leseaufgaben relevant sein.

Was Fragen des Formats und der Aufgabenschwierigkeit betrifft, so hat deutschdidaktische Expertise sich bereits auf die Reflexion zweier Fragenkomplexe eingelassen: „Welche Aufgabenformate und -arten für welche Teilkompetenzen des Textverstehens“ günstig oder ungünstig sind und in welchem Maß und warum bestimmte Gegenstände die Aufgabenschwierigkeit beeinflussen (vgl. Zabka 2006).

- Aufgabengüte

Die drei Gütekriterien für Testaufgaben – „Objektivität“, „Reliabilität“ und „Validität“ – sind in der pädagogischen Diagnostik etabliert. Nicht nur in Handbüchern der Pädagogischen Psychologie (vgl. Wild/Krapp 2001, S. 523-529), sondern auch in deutschdidaktischen Zusammenhängen (z. B. Fix 2006, S. 197f.) werden die drei Kriterien vorgestellt und kommentiert.

Die Aufgabendiskussion der 70er-Jahre hatte vor allem das Kriterium der Auswertungsobjektivität im Blick, die bei offenen Aufgaben und Aufgaben zur Sprachproduktion (Sprechen und Schreiben) nach wie vor eine beträchtliche Herausforderung darstellt und ihrerseits mit zunehmender Komplexität einer Aufgabe wiederum wächst. Demgegenüber wären gegenwärtig – angesichts der Vielzahl ländereigener Vergleichsarbeiten – auch Fragen der Validität (prüft die Aufgabe, was sie prüfen soll?) und der Reliabilität (wie verlässlich ist z. B. eine grammatische Fähigkeit erfasst, wenn zur Überprüfung lediglich ein „Richtig-falsch-Item“ eingesetzt wurde?) zu prüfen. Denn die Einschätzung der Güte von Testaufgaben ist nicht nur eine psychometrische Angelegenheit, sondern verlangt v. a. hinsichtlich der Inhaltsvalidität fachdidaktische und fachwissenschaftliche Expertise. Man muss kennen und verstehen, was Überprüfungsgegenstand ist, wenn man sicher gehen will, dass die Aufgabe das Intendierte misst – also z. B. die Fähigkeit zur Sicherung des inhaltlichen Zusammenhangs eines Textes und nicht nur die Fähigkeit zur Ermittlung einzelner Informationen. Letzteres zeigt sich v. a. in Beiträgen von Kämper-van den Boogaart/Pieper in diesem Band.

Halten wir also fest: Deutschdidaktik ist unverzichtbar, wenn festzulegen gilt, was anlässlich einer Aufgabe erworben oder überprüft werden kann. Im transdisziplinären Verbund der Aufgabenreflexion obliegt ihr vor allem das, was *nur sie* tun kann, nämlich Aufgabenmerkmale strikt gegenstandsbezogen zu prüfen.

Aus der Perspektive ihrer spezifischen Domänen (Sprechen und Schreiben, Zuhören und Lesen) wird sie Wissensvoraussetzungen dimensionieren, Aufgabenformate und -arten diskutieren und schwierigkeitsgenerierende Potenziale der Gegenstände diagnostizieren. Aus der Domänenspezifität der Aufgabenanforderungen resultiert das Bewusstsein des – künftig zu gewinnenden – Zusammenhangs von Aufgaben und Kompetenzmodellen.

- Domänenspezifität der Aufgaben

Auch wenn Deutschdidaktik als Disziplin darauf gründet, dass Kompetenzen in erster Linie domänenspezifisch sind (vgl. Ossner 2006), so richtet sich die Aufmerksamkeit gegenwärtig auf die enge Beziehung zwischen Aufgaben und den Modellierungen der einzelnen Domänen, sprich Lernbereiche. So resümiert Willenberg, „dass es keine Allroundfähigkeit für die einzelnen Gebiete des Deutschunterrichts gibt, sondern dass jeder Teilbereich eine eigene inhärente Stufung entwickeln muss, die aus den zugrunde liegenden Theorien abgeleitet wird“ (Willenberg 2007, S. 14). Das bedeutet, dass die einzelnen Domänen (die KMK-Bildungsstandards sprechen von Kompetenzbereichen) des Deutschunterrichts zum didaktischen Gebrauch einer *inhaltlich prägnanten Strukturierung* bedürfen. Nur in diesem Sinn gut strukturierte Domänen erlauben präzise gegenstandsbezogene Anforderungsbeschreibungen. Für den Bereich des Leseverstehens haben PISA, IGLU und DESI Wesentliches geleistet – die beiden ersten allerdings weitgehend ohne deutschdidaktische Beteiligung. Die Unterscheidungen zum einen von Decodieren und Verstehen und zum anderen – innerhalb des Verstehens – zwischen (1) dem Ermitteln einzelner Informationen, (2) der Sicherung des inhaltlichen Textzusammenhangs und (3) der Beurteilung von Textinformation mithilfe außertextlichen Wissens wurden zum Teil in die KMK-

Bildungsstandards übernommen. Diese Unterscheidungen sind deshalb von hoher heuristischer Qualität, weil sie die Zuordnung von Aufgaben und Fähigkeiten ermöglichen und zugleich transparent machen, dass Lesekompetenz unterschiedliche – wenn auch vielfach vernetzte – Facetten hat.

2.3. Aufgabengenerierung und -codierung

Die Entwicklung von Lernaufgaben wird in erster Linie von Deutschlehrkräften betrieben, sei es für den kleinen Zirkel ihrer Lerngruppen, sei es im großen Rahmen der Lehrmittelentwicklung (vgl. Schäfers 2006). Hier gilt, was in Kapitel 1.1. zur unterrichtlichen Praxis gesagt wurde, und es sei nachdrücklich auf Lechner (2007) verwiesen, der die Aufgabe ins Zentrum didaktischer Reflexion stellt.

Differenzierter gestaltet sich die Lage im Bereich der Leistungsaufgaben. Aufgaben für Zentrale Prüfungen werden einschließlich der entsprechenden Kodieranleitungen (Lösungserwartungen) von dafür durch die Bildungsadministration beauftragten Deutschlehrkräften entwickelt – meist ohne dass diese eine entsprechende Schulung erhalten hätten. Das gerät nicht zuletzt deshalb zu einem Problem, weil offene und zugleich komplexe Aufgaben höchste Anforderungen an Kodieranleitungen stellen. Erschwerend kommt hinzu, dass Aufgaben für Zentrale Prüfungen aus Geheimhaltungsgründen nicht erprobt oder pilotiert werden dürfen, sodass die Gefahr einer Diskrepanz von erwarteten und tatsächlichen Lösungen nicht unerheblich ist.

Anders steht es im Bereich der Testaufgaben. Deren Entwicklung ist psychometrisch geleitet (IQB, IQSH) und unterwirft die von Deutschlehrkräften auf der Basis der KMK-Bildungsstandards generierten domänenspezifischen Aufgaben (Lesen, Schreiben, Zuhören) einer strikten Pilotierungs- und Bewertungsprozedur.

3. Strittiges

Fragt man, was in diesem Forschungsfeld aus deutschdidaktischer Perspektive problematisch erscheint, dann werden vor allem drei Problembereiche sichtbar: der Bereich der Aufgabenarten und -formate, der Bereich der Kompetenzmodelle und das zentrale Ziel des Deutschunterrichts. Deutschdidaktische Kritik an Aufgaben und Kompetenzmodellen richtet sich eher am Rande gegen die aktuelle deutsche Bildungspolitik in toto – deren ökonomische Grundierung durchaus transparent ist; vielmehr sind die im Folgenden exemplarisch aufgeführten deutschdidaktischen Einwände domänen- und gegenstandsspezifisch fundiert und treffen jeweils einen speziellen Träger des Aufgabendiskurses.

3.1. Aufgabenformate und Aufgabenarten

Indem die von Spinner (2004) geäußerten Bedenken gegenüber Multiple-Choice-Aufgaben *Fragen des Aufgabenformats bzw. der Aufgabenart* berühren, beziehen sie zugleich Front gegen die wissenschaftliche Disziplin, die dieses Format – aus guten Gründen – präferiert: gegen die Psychometrie. Aus literaturdidaktischer Perspektive – also strikt domänenbezogen – stellt Spinner fest, dass „solche Aufgabenformate den Verstehensprozess nur indirekt“ abbildeten, denn sie lösten „einen kognitiven

Denkprozess aus, der sich von einer direkten Begegnung mit einem Text unterscheidet“ (Spinner 2004, S. 134).

Im Anschluss daran wäre zu überlegen, ob zur Überprüfung textbezogener Verstehensleistungen insbesondere bei literarischen Texten nicht vielleicht doch offene Formate gefragt sind – mit allen Kodierungsproblemen, die daraus resultieren mögen (vgl. Schneider/Bertschi-Kaufmann 2006). Das gilt sowohl für die Überprüfung globaler Kohärenzbildung als auch für die Sinnkonstruktion bei literarischen Texten. Hier wären Studien erwünscht, die den Einsatz von Multiple-Choice-Aufgaben mit Begründungsauftrag hinsichtlich der getroffenen Wahl evaluieren. Denn bei der Generierung hierarchiehöherer Textverstehensaufgaben zeigen sich zwei Probleme: 1. Distraktoren zu finden, die sowohl plausibel als auch unzutreffend sind, und 2. sicher zu stellen, dass die Distraktoren noch das überprüfen, was sie überprüfen sollen – nämlich „textbezogene Interpretationsleistung“ und nicht die Genauigkeit, mit der das Auswahlangebot gelesen wird. Hier wird nur die interessierte Kooperation beider Disziplinen Gewinne verbuchen lassen. Diskutable Angebote finden sich im Beitrag von Kämper-van den Boogaart/Pieper in diesem Band.

Die strittige Frage, auf welche Weise das Verstehen von – literarischen – Texten am besten überprüft wird, stellt sich jedoch nicht nur von den geschlossenen Multiple-Choice-Formaten her, sondern auch anlässlich der von den Kultusministerien organisierten Aufgaben für die Zentralen Abschlussprüfungen, vor allem im Abitur. Da die traditionellen Prüfungsaufgaben im Zuge ihrer Erschließungs- oder Erörterungsanforderungen zwei zunächst unabhängige Domänen – Lesen und Schreiben – auf das Engste verquicken, steht auch hier die Konstruktvalidität auf dem Spiel. Denn was wird überprüft, wenn es die schriftliche Produktionsleistung ist, die Auskunft über das Textverstehen und dessen Ausprägungsgrad gibt? Deutschlehrkräfte wissen: Wer anspruchsvolle Texte versteht, muss noch kein guter Schreiber sein – und das gilt auch vice versa. Rezeptive und produktive Kompetenzen halten sich keineswegs die Waage. Folglich hätte deutschdidaktische Reflexion im Verbund mit der Bildungsadministration und der Psychometrie empirisch zu ermitteln, ob und ggf. in welchem Maß der Ausweis literarischen Verstehens – und des Textverstehens auf globaler Ebene überhaupt – von der schriftlichen Produktion getrennt werden kann.

Zentrale Prüfungsaufgaben traditionellen Zuschnitts bringen noch ein weiteres Güteproblem auf den Plan. So plausibel und notwendig es erscheint, bei zentralen Prüfungen auch Auswertungsobjektivität zu gewährleisten, so schwierig gestaltet sich das bei offenen und zugleich hochkomplexen Aufgaben. Sowohl die Art als auch der Umfang detaillierter und gewichteter inhaltsbezogener Codieranleitungen erweisen sich als stark optimierungsbedürftig (vgl. Kap. 2.3.), da sie bei der Auswertung zu keiner größeren Interraterobjektivität führen als holistische Beurteilungen anhand traditioneller (implementierter bzw. EPA-gestützter) Kriterien. Die fachdidaktische Erklärung dieses Befunds und Hinweise für die Modifizierung der Codieranleitungen sind als Resultat der Evaluation des ersten Zentralabiturs Deutsch in Nordrhein-Westfalen zu erwarten.

Fazit: Diese beiden domänenspezifisch basierten Frontstellungen gegen Psychometrie auf der einen und die Bildungsadministration auf der anderen Seite sind aus zwei Gründen interessant: Zum einen, weil sie die Pole markieren, zwischen denen Leistungsaufgaben gegenwärtig angesiedelt sind, sodass die Größe dieses Spannungsfelds sichtbar wird.

Zum anderen, weil die deutschdidaktische Offensive zwei Diskursträgern gilt – der Psychometrie und der Bildungsadministration –, die partiell durchaus Diskurskoalitionäre sind. Anders formuliert: Die Bildungsadministration ist in Sachen „Leistungsaufgaben“ kein Monist, sondern verfolgt zwei sehr unterschiedliche Linien – die der Tradition des Deutschunterrichts (mit dem integralen Ziel einer schriftlichen Texterschließungs- oder Texterörterungsleistung) und die Linie der Testkultur (mit dem Ziel objektiver, verlässlicher und gültiger Leistungsmessungen in den Teilbereichen einzelner Domänen).

Unter den Trägern des Aufgabendiskurses hat die Bildungsadministration eine starke Stimme und Gewicht. Deutschdidaktische Expertise wird gebraucht, um die Vielfalt/Vielzahl der Möglichkeiten *zwischen* diesen Polen – innerhalb des Spannungsfelds – zu explorieren.

3.2. Kompetenzmodelle

Ebenso strittig sind Bemühungen, differenzierte, gestufte Kompetenzmodelle zu generieren und entsprechend differenzierende Aufgaben zu entwickeln. Psychometrie und die mit ihr vernetzten Teile der Bildungsadministration brauchen fachdidaktisch-theoretisch fundierte Kompetenzmodelle, denn diese bilden die Voraussetzung für Testmodelle. Aus dieser Perspektive wird auch der Zusammenhang von Aufgaben und Kompetenzmodellen plausibel. Allerdings ist – das zeigt die aktuelle deutschdidaktische Diskussion (Didaktik Deutsch 21/22/23) – die Entwicklung von Kompetenzmodellen ein Großprojekt mit offenem Ausgang. Es erscheint sehr schwierig, komplexe Prozesse wie das Verstehen literarischer Texte oder das Schreiben argumentativer Texte so zu strukturieren, dass sie – weitgehend unabhängig vom konkreten Gegenstand – in aussagekräftigen Kompetenzmodellen organisiert werden können.

Strittig bleibt daher die Frage, in *welchem Maß* sich die deutschdidaktische Aufgabenreflexion in den Dienst der Entwicklung von Kompetenzmodellen stellen soll. Wenn Kompetenzmodelle aus deutschdidaktischer Perspektive jedoch primär Hilfsfunktion zu übernehmen hätten, um über die Anforderungen und Voraussetzungen einer Aufgabe verlässlich Auskunft zu geben, dann erscheint dieses Unternehmen durchaus als lohnend. Aufgaben im Kontext domänenspezifischer Kompetenzmodelle zu entwickeln und zu verorten – das bleibt dennoch auf längere Sicht ein Desiderat.

3.3. Ziel des Deutschunterrichts

Strittig ist darüber hinaus, wie hoch die „Gefahr einer einseitigen Fixierung des Unterrichts auf das Testbare und die Testsituation“ (Kammler 2006, S. 20) einzuschätzen ist. Wer die Befürchtung teilt, dass die Lösung von Testaufgaben zum Zentrum

des Deutschunterrichts wird, setzt sich mit den Bildungsmedien auseinander. Da diese sehr schnell auf die neuen Konzepte und Bedürfnisse reagiert haben, erscheint die Kritik an der Fülle von „Trainings-“ und „Mach dich fit!-Angeboten“ (Spinner 2005, S. 10) durchaus berechtigt. Die Befürchtung, dass nicht nur „Situationen des Staunens“ und „des selbstvergessenen Lesens“ (Spinner 2005) auf der Strecke bleiben, sondern auch Aufgaben, die individuelle Lösungswege und spekulative Lösungen zulassen, wird von vielen Lehrkräften geteilt. Kritik gilt in diesem Zusammenhang vor allem der Vermarktung von Trainingsaufgaben ganz unterschiedlicher Qualität ohne Reflexion auf den Bildungswert.

Dieser Gefahr ist zu begegnen. Denn guter Unterricht braucht sowohl Einübung von Routinen als auch Exploration von Neuem, und das auf vielfältigen Wegen. Vernünftige Lehrkräfte wissen das und handeln entsprechend. Gegebenenfalls hätte die Bildungsadministration die notwendige Balance einzufordern. Denn die Einführung von Leistungstests dient in erster Linie der Sicherung von Basiskompetenzen. Gesellschaftliche Partizipation verlangt aber vieles mehr.

Literatur

- Artelt, Cordula/Stanat, Petra/Schneider, Wolfgang/Schiefele, Ulrich/Lehmann, Rainer (2004): Die PISA-Studie zur Lesekompetenz. Überblick und weiterführende Analysen. In: U. Schiefele/C. Artelt/W. Schneider/P. Stanat (Hgg.): Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS, S. 139-168.
- Artelt, Cordula/McElvany, Nele/Christmann, Ursula/Richter, Tobias/Groeben, Norbert/Köster, Juliane/Schneider, Wolfgang/Stanat, Petra/Ostermeier, Christian/Schiefele, Ulrich/Valtin, Renate/Ring, Klaus (2005): Förderung von Lesekompetenz. Berlin: BMBF (Bildungsreform Bonn 17), S. 12.
- Bremerich-Vos, Albert (2003): Bildungsstandards, Kompetenzstufen, Kernlehrpläne, Parallel- und Vergleichsarbeiten. In: Deutschunterricht 56, Heft 5, S. 4-10.
- Fix, Martin (2006): Texte schreiben. Schreibprozesse im Deutschunterricht. Paderborn u. a.: Schöningh, S. 188-207.
- Funk, Hermann (2006): Aufgabenorientierung in Lehrwerk und Unterricht – das Problem der Theorie mit der Vielfalt der Praxis. In: K.-R. Bausch/E. Burwitz-Melzer/F. G. Königs/H.-J. Krumm (Hgg.): Aufgabenorientierung als Aufgabe. Arbeitspapiere der 26. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts. Tübingen: Narr (Gießener Beiträge zur Fremdsprachendidaktik), S. 52-61.
- Haas, Gerhard/Rupp, Gerhard/Waldmann, Günter (1989): Produktionsorientierter Umgang mit Literatur in der Schule. Abschließende Bemerkung zur Theoriediskussion und Anregung für die Praxis (nebst einem geharnischten Brief). In: Praxis Deutsch 98, S. 6-13.
- Häusermann, Ulrich/Piepho, Hans-Eberhard (1996): Aufgaben-Handbuch. Deutsch als Fremdsprache. Abriß einer Aufgaben- und Übungstypologie. München: Iudicium.
- Ingenkamp, Karlheinz [Hg.] (1971): Test in der Schulpraxis. Eine Einführung in Aufgabenstellung, Beurteilung und Anwendung von Tests. Weinheim u. a.: Beltz.
- Jenkins, James J. (1979): Four points to remember. A tetrahedral model and memory experiments. In: L. S. Cermak/F. I. M. Craik (Hgg.): Levels of processing in human memory. Hillsdale: Erlbaum, S. 429-446.
- Kammler, Clemens (2006): Literarische Kompetenzen – Standards im Literaturunterricht. Anmerkungen zum Diskussionsstand. In: C. Kammler (Hg.): Literarische Kompetenzen – Standards im Literaturunterricht. Seelze: Kallmeyer, S. 7-22.
- Köster, Juliane (2003): Die Profilierung einer Differenz. Aufgaben zum Textverstehen in Lern- und Leistungssituationen. In: Deutschunterricht 56, Heft 5, S. 19-25.
- Köster, Juliane (2005): Evaluation von Kompetenzen im Deutschunterricht – neues Etikett oder bildungspolitische Wende? In: H. Rösch (Hg.): Kompetenzen im Deutschunterricht. Frankfurt a. M.: Lang, S. 53-71.
- Köster, Juliane (2006a): Unterschiedliche Prüfungsmuster im deutschen Schulwesen. In: K.-M. Kodalle (Hg.): Der geprüfte Mensch. Über Sinn und Unsinn des Prüfungswezens. Würzburg: Königshausen&Neumann, S. 25-36.

- Köster, Juliane (2006b): Das Deutschabitur in Zeiten von Bildungsstandards – Vergleichbarkeit der Prüfungsleistungen und ihrer Bewertung. In: *Didaktik Deutsch* 21, S. 78-90.
- Lechner, Manfred (2007): Unterrichten heißt: Aufgaben stellen. Befunde und Konsequenzen. In: *Seminar* 2, S. 122-133.
- Legutke, Michael K. (2006): Aufgabe – Projekt – Szenario. Über die großen Perspektiven und die kleinen Schritte. In: K.-R. Bausch/E. Burwitz-Melzer/F. G. Königs/ H.-J. Krumm (Hgg.): *Aufgabenorientierung als Aufgabe*. Arbeitspapiere der 26. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts. Tübingen: Narr (Giebener Beiträge zur Fremdsprachendidaktik), S. 140-148.
- Lindauer, Thomas/Schneider, Hansjakob (2007): Lesekompetenz ermitteln. Aufgaben im Unterricht. In: A. Bertschi-Kaufmann (Hg.): *Lesekompetenz – Leseleistung – Leseförderung*. Grundlagen, Modelle und Materialien. Zug: Klett und Balmer (= *Lehren lernen* 1), S. 109-125.
- Ossner, Jakob (2006): Kompetenzen und Kompetenzmodelle im Deutschunterricht. In: *Didaktik Deutsch* 21, S. 5-19.
- Schäfers, Stefanie (2006): Aufgabenstellungen im Deutschunterricht. Eine Anleitung zur Formulierung verständlicher schriftlicher Aufgaben in der gymnasialen Oberstufe aus Sicht der Sprachwissenschaften. Münster: LIT.
- Schneider, Hansjakob/Bertschi-Kaufmann, Andrea (2006): Lese- und Schreibkompetenzen fördern. Erste Ergebnisse einer Interventionsstudie. In: *Didaktik Deutsch* 20, S. 30-51.
- Schneider Hansjakob/Lindauer, Thomas (2007): Lesekompetenz ermitteln: Tests. In: A. Bertschi-Kaufmann (Hg.): *Lesekompetenz – Leseleistung – Leseförderung*. Grundlagen, Modelle und Materialien. Zug: Klett und Balmer (= *Lehren lernen* 1). S. 126-139.
- Schnotz, Wolfgang/Dutke, Stephan (2004): Kognitionspsychologische Grundlagen der Lesekompetenz: Mehrebenenverarbeitung anhand multipler Informationsquellen. In: U. Schiefele/C. Artelt/W. Schneider/P. Stanat (Hgg.): *Struktur, Entwicklung und Förderung von Lesekompetenz*. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS, S. 61-99.
- Schweitzer, Kirstin (2007): Der Schwierigkeitsgrad von Textverstehensaufgaben. Ein Beitrag zur Differenzierung und Präzisierung von Aufgabenbeschreibungen. Frankfurt a. M.: Lang.
- Spinner, Kaspar H. (2004): Lesekompetenz in der Schule. In: U. Schiefele/C. Artelt/W. Schneider/P. Stanat (Hgg.): *Struktur, Entwicklung und Förderung von Lesekompetenz*. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS, S. 125-138.
- Spinner, Kaspar H. (2005): Der standardisierte Schüler. In: *Didaktik Deutsch* 18, S. 4-13.
- Weinert, Franz E. (1999): Die fünf Irrtümer der Schulreformer. In: *Psychologie heute* 7, S. 28-34.
- Wild, Klaus-Peter/Krapp, Andreas (⁴2001): Pädagogisch-psychologische Diagnostik. In: A. Krapp/B. Weidemann (Hgg.): *Pädagogische Psychologie*. Ein Lehrbuch. Weinheim: Beltz, S. 513-564.
- Willenberg, Heiner (2007): Die Kompetenztheorie für den Deutschunterricht baut noch kein Haus. In: *Didaktik Deutsch* 22, S. 13-22.

- Winkler, Iris (2005): Zur Beziehung von Unterrichtsmaterial, -gestaltung und -erfolg: Drei Aufgaben zu Georg Brittings „Brudermord im Altwasser“ im Praxistest. In: J. Stückrath/R. Strobel (Hgg.): Deutschunterricht empirisch. Beiträge zur Überprüfbarkeit von Lernfortschritten im Sprach-, Literatur- und Medienunterricht. Baltmannsweiler: Schneider Hohengehren, S. 177-196.
- Zabka, Thomas (2006): Typische Operationen literarischen Verstehens. Zu Martin Luther „Vom Raben und Fuchs“ (5./6. Schuljahr). In: C. Kammler (Hg.): Literarische Kompetenzen – Standards im Literaturunterricht. Seelze: Kallmeyer, S. 80-101.