



Winfried Ulrich (Hrsg.)

Deutschunterricht in Theorie und Praxis (DTP)

11/3

Lese- und Literaturunterricht

Teil 3: - Erfolgskontrollen und
Leistungsmessung
- Exemplarische Unterrichtsmodelle

Michael Kämper-van den Boogaart
Kaspar H. Spinner
(Hrsg.)



Deutschunterricht in Theorie und Praxis

DTP

**Handbuch zur Didaktik der deutschen Sprache
und Literatur in elf Bänden**

herausgegeben von

Winfried Ulrich

Band 11 _{/3}



Schneider Verlag Hohengehren GmbH

Lese- und Literaturunterricht

Teil 3
Erfolgskontrollen und Leistungsmessung
Exemplarische Unterrichtsmodelle

3., stark überarbeitete Auflage

herausgegeben von
Michael Kämper-van den Boogaart
und
Kaspar H. Spinner



Schneider Verlag Hohengehren GmbH

Umschlag: PE-Mediendesign, Elke Boscher,
88521 Ertingen

Gedruckt auf umweltfreundlichem Papier (chlor- und säurefrei hergestellt).

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

ISBN: **978-3-8340-1933-2**

ISBN: 978-3-8340-1950-9 (→ Band 11.1 bis 11.3)

Schneider Verlag Hohengehren, Wilhelmstr. 13, D-73666 Baltmannsweiler

Homepage: www.paedagogik.de

Das Werk und seine Teile sind urheberrechtlich geschützt. Jede Verwertung in anderen als den gesetzlich zugelassenen Fällen bedarf der vorherigen schriftlichen Einwilligung des Verlages. Hinweis zu § 52 a UrhG: Weder das Werk noch seine Teile dürfen ohne vorherige schriftliche Einwilligung des Verlages öffentlich zugänglich gemacht werden. Dies gilt auch bei einer entsprechenden Nutzung für Unterrichtszwecke!

Inhaltsverzeichnis

Vorwort des Herausgebers des Handbuchs „Deutschunterricht in Theorie und Praxis“	VII
---	-----

Vorwort der Herausgeber der drei lese- und literaturdidaktischen Bände	XI
---	----

F Erfolgskontrollen und Leistungsmessung

F 1	Aufgabentypen für Erfolgskontrollen und Leistungsmessung im Literaturunterricht <i>von Juliane Köster</i>	3
F 2	Psychometrische Erfassung von Lesekompetenz in internationalen und nationalen Leistungsstudien <i>von Andreas Voss/Inge Blatt</i>	29
F 3	Texte über Texte als Formate schriftlicher Leistungsüberprüfung: Nacherzählung, Inhaltsangabe, Analyse, Interpretation und benachbarte Aufgaben <i>von Thomas Zabka</i>	63
F 4	Bewertung bei handlungs- und produktionsorientierten Schreibaufgaben und beim literarischen Schreiben <i>von Ulf Abraham</i>	93
F 5	Bewerten mündlicher Leistungen im Literaturunterricht <i>von Kerstin Hillegeist</i>	120

G Exemplarische Unterrichtsmodelle

G 1	Lesetechniken erlernen, Lesestrategien entwickeln im Unterricht <i>von Christoph Bräuer</i>	155
G 2	Analyse und Interpretation literarischer Texte – exemplarische Unterrichtsmodelle <i>von Dieter Wrobel</i>	201
G 3	Zum Umgang mit Sachtexten im Deutschunterricht: thematisch und integrativ <i>von Astrid Müller</i>	237
G 4	Literaturgeschichte im Unterricht als Kulturgeschichte <i>von Karlheinz Fingerhut</i>	256

G5	Handlungs- und Produktionsorientierung – ein Modell für die Grundschule: Jutta Richter: „Hinter dem Bahnhof liegt das Meer“ (Klasse 4) von <i>Ingrid Hintz</i>	298
G6	Handlungs- und Produktionsorientierung – ein Modell für die 9.–11. Jahrgangsstufe: Charles Simmons „Salzwasser“ von <i>Monica Gross</i>	317
G7	Szenische Interpretation – ein Unterrichtsmodell für die Primarstufe: Die Sprache der Gefühle von <i>Christina Dieterle/Uta-Ena Iaconis</i>	337
G8	Szenische Interpretation – ein Unterrichtsmodell für die Sekundarstufe I: Heinrich Böll „Die Waage der Baleks“ (8./9. Klasse) von <i>Christina Dieterle/Uta-Ena Iaconis</i>	352
Sachwortregister		373

Vorwort des Herausgebers des Handbuchs

Zur Konzeption des Handbuches DTP

Der Deutschunterricht ist das zentrale Unterrichtsfach, das Kernfach im allgemeinbildenden Schulwesen Deutschlands mit der Aufgabe, nicht nur die grundlegenden Kulturtechniken des Lesens und Schreibens zu vermitteln, sondern die an Sprache und Texte gebundenen kognitiven, kommunikativen und kreativen Fähigkeiten der heranwachsenden Generation zu fördern. Dabei geht es um „Schlüsselqualifikationen“, die zur Bewältigung vieler Situationen im privaten und beruflichen Alltag benötigt werden, sowie um Kompetenzen, die zur Teilnahme am kulturellen Leben der Gesellschaft befähigen.

Sprache ist von zentraler Bedeutung für das menschliche Leben. Erst seine Sprache macht den Menschen zu dem, der er ist. Seine Sprachkompetenz befähigt ihn zum Handeln. Sie hilft beim Erwerb von Weltwissen, bei der Begriffsbildung und bei Denkopoperationen; sie erlaubt den zwischenmenschlichen Austausch von Gedanken und Gefühlen, rezeptiv beim Zuhören und Lesen, produktiv beim Sprechen und Schreiben. Keine andere Fähigkeit ist für die Persönlichkeitsentfaltung und die Entwicklung zu einem mündigen Glied der Gesellschaft von ähnlich herausragender Bedeutung.

Sprachkompetenz wird von früher Kindheit an erworben und erweitert, zunächst auf „natürliche“ Weise in der Familie, dann in Vorschule und Schule systematisch und zielorientiert angeleitet in Lehr- und Lernprozessen, vor allem im Deutschunterricht. Auf die Unterrichtspraxis in den Schulen und die dort erzielten Ergebnisse richten sich deshalb häufig sowohl die kritischen Blicke der Öffentlichkeit wie auch die hoffnungsvollen Erwartungen von Eltern und Bildungspolitikern.

Welche Aufgaben soll der Deutschunterricht haben? Wie soll Deutschunterricht erteilt werden? Eine Antwort auf diese Fragen darf von der „Fachdidaktik Deutsch“ erwartet werden, der Theorie des Deutschunterrichts. Sie ist eine in Forschung und Lehre an den Universitäten und Pädagogischen Hochschulen vertretene wissenschaftliche Disziplin. Als solche ist sie relativ jung und bislang noch nicht voll entwickelt. Sie hat es gegenwärtig nicht leicht. Einerseits muss sie sich an den Hochschulen im Kreis etablierter Disziplinen behaupten und mit vorzeigbaren, kritisch diskutierten Forschungsergebnissen um Anerkennung kämpfen. Andererseits darf sie sich dabei nicht in wirklichkeitsferne Theorie verirren. Sie muss die gegebene Unterrichtspraxis, die konkreten Lehr- und Lernprozesse im Alltag der Schulen und sonstigen Bildungseinrichtungen im Auge behalten und sich innovativ auf diese beziehen. Erst in der unterrichtspraktischen Erprobung erweist sich, ob die theoretischen Konstrukte umsetzbar sind oder modifiziert werden müssen. Eine (möglichst

institutionalisierte) ständige Kommunikation zwischen Forschern und Praktikern ist notwendig, der wechselseitige Austausch von Impulsen und Erfahrungen. Praxisorientierte Theorie an den Hochschulen und theoriegeleitete Praxis in den Schulen: Das muss als übergeordnetes Ziel gelten.

Diesem Ziel ist das vorliegende dreizehnbändige Handbuch „Deutschunterricht in Theorie und Praxis“ (DTP) verpflichtet. Mit ihm wird erstmals der Versuch unternommen, ein Schulfach einer in diesem Ausmaß noch nie durchgeführten wissenschaftlichen Bestandsaufnahme zu unterziehen und aus dieser Bestandsaufnahme schlüssige Vorschläge für eine zukünftige Weiterentwicklung abzuleiten. Die Theorie des Deutschunterrichts in der klassischen Ausformung von Sprachunterricht und Literaturunterricht unterzieht sich damit einer Selbstevaluation, besinnt sich auf ihre Entwicklung und ihren gegenwärtigen Zustand, artikuliert ihr Selbstverständnis und beschreibt antizipierend ihre Zukunft. Dabei bekennt sie sich zu den in den letzten Jahren/Jahrzehnten erfolgten Erweiterungen ihres Zuständigkeitsbereichs: So sind z. B. Deutsch als Zweitsprache, Deutsch als Fremdsprache, Mediendidaktik und frühkindliche Sprachförderung in den Fokus der Deutschdidaktik gerückt.

Die Komplexität des Gegenstands „Fachdidaktik Deutsch“ und die Schwierigkeiten seiner Beschreibung zeigen sich auch darin, dass es in vielen Aspekten Überschneidungen zu anderen Wissenschaften gibt, die sich mit der Ontogenese, der Erziehung, der Psyche von Kindern, der Sprache als System und Gebrauch, der Literatur, den Kommunikationsmedien usw. beschäftigen. Interdisziplinäre Zusammenarbeit mit gegenseitiger Befruchtung ist notwendig, z. B. mit den unmittelbaren Bezugswissenschaften Linguistik und Literaturwissenschaft, aber auch mit Schulpädagogik, Lernpsychologie, den Kognitionswissenschaften.

Außer der Zusammenarbeit mit anderen Wissenschaftsdisziplinen ist die „Fachdidaktik Deutsch“ auf empirische Unterrichtsforschung angewiesen. Eine solche ist bislang erst in Ansätzen vorhanden, muss aber für die Zukunft in größerem Umfang gefordert und durchgeführt werden. Das gilt besonders, wenn aktuell Schulen und Lehrkräften einerseits viel mehr Gestaltungsfreiheit eingeräumt wird als bisher, andererseits der Erfolg ihrer Arbeit mit Hilfe fester Bildungsstandards überprüft werden soll. Beides ist erklärter Wille der Bildungspolitiker.

Das Handbuch DTP beschreibt also den Deutschunterricht nach Zielsetzungen, Unterrichtsmethoden, Unterrichtsmitteln, Lernergebnissen. Eine systematische Beschreibung kommt nicht umhin, das Schulfach als eine Zusammensetzung aus verschiedenen Ziel- und Inhaltsbereichen anzusehen. Eine solche Aufteilung stellt in gewisser Weise eine künstliche Trennung von Zusammengehörendem dar, denn natürlich sind die Bereiche nicht autonom, sondern bereits in der Theorie vielfältig miteinander verflochten, erst recht in der Unterrichtspraxis integrativ zu behandeln. Im Bewusstsein dieser Spannung orientiert sich das Gesamtwerk DTP in seiner Gliederung an den Lern- und Arbeitsbereichen, die sich aus der Curriculardiskus-

sion der letzten Jahrzehnte ergeben haben, und ergänzt sie um weitere für die Deutschdidaktik relevante Bereiche. So ergeben sich insgesamt elf Themenbereiche (verteilt auf dreizehn Bände bzw. Teilbände), die in der Verantwortung von jeweiligen Bandherausgebern liegen:

- 1 Deutsche Sprache in Kindergarten und Vorschule
Prof. Dr. Herbert Günther, Universität Koblenz-Landau
- 2 Schriftsprach- und Orthographieerwerb: Erstlesen, Erstschreiben
Prof. Dr. Christa Röber-Siekmeyer, Pädagogische Hochschule Freiburg
- 3 Mündliche Kommunikation und Gesprächsdidaktik
Prof. Dr. Michael Becker-Mrotzek, Universität Köln
- 4 Schriftlicher Sprachgebrauch/Texte verfassen
Prof. Dr. Helmuth Feilke, Universität Gießen
Prof. Dr. Thorsten Pohl, Universität Oldenburg
- 5 Weiterführender Orthographieerwerb
Prof. Dr. Ursula Bredel, Universität Köln
- 6 Sprachreflexion und Grammatikunterricht
Prof. Dr. Hildegard Gornik, Universität Hildesheim
- 7 Wortschatzarbeit
Prof. Dr. Inge Pohl, Universität Koblenz-Landau
Prof. Dr. Winfried Ulrich, Universität Kiel
- 8 Digitale Medien im Deutschunterricht
Prof. Dr. Volker Frederking, Universität Erlangen-Nürnberg
Prof. Dr. Thomas Möbius, Universität Gießen
Axel Krommer, Universität Erlangen-Nürnberg
- 9 Deutsch als Zweitsprache
Prof. Dr. Ingelore Oomen-Welke, Pädagogische Hochschule Freiburg
Prof. Dr. Bernt Ahrenholz, Universität Jena
- 10 Deutsch als Fremdsprache
Prof. Dr. Ingelore Oomen-Welke, Pädagogische Hochschule Freiburg
Prof. Dr. Bernt Ahrenholz, Universität Jena
- 11–13 Lese- und Literaturunterricht, Bände 1–3
Prof. Dr. Kaspar Spinner, Universität Augsburg
Prof. Dr. Michael Kämper-van den Boogaart, Humboldt-Universität Berlin

Als Projektleiter und Gesamtherausgeber des Werkes zeichnet Prof. Dr. Dr. h. c. mult. Winfried Ulrich, Germanistisches Seminar der Christian-Albrechts-Universität zu Kiel, verantwortlich.

Alle Teilbände fügen sich in eine gemeinsame Gesamtkonzeption ein, andererseits sind sie auch als in sich geschlossenen Abhandlungen zu verstehen. Inhaltliche

Überlappungen der Einzelbände sind nicht ganz zu vermeiden, bis zu einem gewissen Grad sogar wünschenswert. So gibt es z. B. einen speziellen Band für Deutsch als Zweitsprache; unabhängig davon muss die Berücksichtigung der unterschiedlichen Herkunftssprachen von Schülern mit Migrationshintergrund auch in anderen Bänden erfolgen. Obgleich sich die einzelnen Bände marginal aufgrund des jeweiligen Standes der Forschung in Aufbau und Umfang unterscheiden, orientieren sich alle Autorinnen und Autoren grundsätzlich an folgender Binnengliederung:

- 1 Geschichte und Entwicklung des Teilbereichs des Deutschunterrichts
- 2 Konzeptionelle und empirische Grundlagen
- 3 Kompetenzbereiche, Unterrichtsziele
- 4 Methoden und Medieneinsatz
- 5 Gegenwärtiger Stand empirischer Unterrichtsforschung
- 6 Erfolgskontrollen und Leistungsmessung
- 7 Exemplarische Unterrichtsmodelle

Der Herausgeber des Gesamtwerks und die Herausgeberinnen und Herausgeber der Einzelbände hoffen, dass von dieser Veröffentlichung über viele Jahre hin fruchtbare Impulse ausgehen werden: erstens auf die schulpolitischen Diskussionen und Entscheidungen (einschließlich Lehrerbildung und Lehrplanentwicklung in den einzelnen Bundesländern), zweitens auf die fachdidaktische Forschung und Lehre an Universitäten und Pädagogischen Hochschulen, drittens auf eine stärker praxisorientierte Lehrerbildung in den Studienseminaren und viertens schließlich auf die konkrete Unterrichtsgestaltung in den Schulen.

Möge im Umfeld der Wissenschaften die mit DTP vorgelegte Strukturierung der Theorie des Deutschunterrichts zusammen mit den unterbreiteten Vorschlägen zu vertiefter Forschung in ihren Gegenstandsbereichen einer weiteren Präzisierung und Profilierung der „Fachdidaktik Deutsch“ dienlich sein!

Stampe, im Sommer 2008

Winfried Ulrich

Die Konzeption des Handbuches hat sich nicht verändert. Das Vorwort des Herausgebers ist auch noch nach zehn Jahren aktuell.

Stampe, im Sommer 2018

Winfried Ulrich

Vorwort der Herausgeber der drei lese- und literaturdidaktischen Bände

Das Lesen ist mit Beginn des 21. Jahrhunderts zu einem zentralen bildungspolitischen Thema geworden. Lesefähigkeit wird, vor allem seit der ersten PISA-Studie, als eine Schlüsselqualifikation betrachtet, die für Beruf und privates Leben in der zivilisatorisch entwickelten Gesellschaft der Gegenwart grundlegend ist. Entsprechend hat die Forschung zur Lesedidaktik in den letzten Jahren international eine enorme Steigerung erfahren, deren Auswirkungen auf die Schulpraxis zunehmend sichtbar werden. Dabei ergibt sich ein Spannungsverhältnis zum bisher in den deutschsprachigen Ländern vorherrschenden Verständnis von Literaturunterricht, weil der international gebräuchliche Begriff der „Literacy“ mehr am Umgang mit nichtliterarischen als mit literarischen Texten orientiert ist. Dies hat in der deutschdidaktischen Diskussion nicht nur zu verstärkter Forschung im Bereich der Lesekompetenz geführt, sondern gerade auch ein neues Nachdenken über die Rolle, die der Literatur im Unterricht zukommen soll, ausgelöst. Die vorliegenden drei Bände sind entsprechend mit „Lese- und Literaturunterricht“ betitelt und greifen mit dieser Doppelformulierung die beiden hauptsächlichen Schwerpunktsetzungen in der didaktischen Diskussion auf. Die Tatsache, dass „Lesen“ und „Literatur“ kategorial eigentlich nicht auf der gleichen Ebene liegen (wie man kritisch vermerken könnte), soll dabei als Hinweis auf die Überlappungen verstanden werden. Die einzelnen Beiträge in den Bänden beziehen sich zum Teil schwerpunktmäßig nur auf die Lese- oder die Literaturdidaktik (so gibt es einen Beitrag zur Modellierung der Lesekompetenz und einen anderen zur literarischen Rezeptionskompetenz), teilweise aber auch auf beide Aspekte (z. B. die schulstufenspezifischen Beiträge in Teil C). Darin spiegelt sich die Tatsache, dass in der gegenwärtigen didaktischen Diskussion und Forschung Lese- und Literaturunterricht mal getrennt und mal integriert zum Thema gemacht werden.

Die drei Bände zum Lese- und Literaturunterricht bilden eine Einheit; die Grundstruktur, die für alle DTP-Bände gilt, erstreckt sich hier entsprechend auf drei Bände. Band 11,1 gilt der Geschichte des Lese- und Literaturunterrichts und den konzeptionellen und empirischen Grundlagen. Als „Grundlagen“ haben wir die Lesesozialisation und -psychologie, die Literatur- und Texttheorie und die Kompetenzorientierung ausgewählt. Damit soll nicht gesagt sein, dass sich die theoretischen Grundlagen, die für die Lese- und Literaturdidaktik wichtig sind, darauf beschränken – Bildungstheorie und Entwicklungspsychologie wären z. B. ebenfalls zu nennen. Deren Bedeutung wird hier nicht in eigenen Beiträgen erörtert, entsprechende Zusammenhänge werden aber innerhalb anderer Beiträge angesprochen.

In zweiten Band (11,2) finden sich Beiträge zu den Kompetenzen und Unterrichtszielen, die für die einzelnen Schulstufen leitend sind, ergänzt durch einen Beitrag zu den fachspezifischen Lehrerkompetenzen (Teil C). Ein zweiter Teil des Bandes gilt den Unterrichtsmethoden im Lese- und Literaturunterricht und den Unterrichtsmaterialien (Teil D). Schließlich wird in Teil E im selben Band über den Stand der empirischen Unterrichtsforschung informiert; zur Konkretisierung werden auch zwei Forschungsprojekte ausführlicher vorgestellt.

Der dritte Band (11,3) gilt in einem ersten Teil F der Leistungsbeurteilung und –messung. Da durch die großen Evaluationsstudien und die Bildungsstandards die Frage der Leistungsmessung bildungspolitisch und fachwissenschaftlich derzeit in noch nie dagewesener Intensität diskutiert wird, werden entsprechende Fragestellungen theoretisch und praxisbezogen in fünf Beiträgen des Bandes ausführlich dargestellt. Unterrichtsmodelle schließen den Band ab (Teil G); sie sind jeweils auf bestimmte methodische Vorgehensweisen fokussiert. Es versteht sich, dass verschiedene methodische Zugänge in der Praxis nicht getrennt in einzelnen Unterrichtseinheiten erfolgen (sollen) – so hat z. B. ein handlungs- und produktionsorientierter Literaturunterricht auch mit Interpretieren zu tun und szenische Interpretation kann der Auseinandersetzung mit Literaturgeschichte dienen. Dies wird in den Unterrichtsmodellen, auch wenn sie einem bestimmten methodischen Verfahren gelten, auch deutlich gemacht.

Da die drei lese- und literaturdidaktischen Bände im Kontext des gesamten DTP-Werkes zu sehen sind, werden einige Fragestellungen, die man eventuell in ihnen erwartet, nur angeschnitten. Das gilt vor allem für den Umgang mit den digitalen Medien. Da Kindern und Jugendlichen Literatur überwiegend im Medienverbund begegnet und ihre alltägliche Lesepraxis zu einem großen Teil am Bildschirm und mit dem Handy erfolgt, muss Lese- und Literaturunterricht auch zu diesen Medien einen Bezug herstellen. Dieser Aspekt wird in den vorliegenden literatur- und lese-didaktischen Bänden zwar angesprochen, ausführlich entfaltet wird er jedoch in DTP 8, dem Band zu den digitalen Medien im Deutschunterricht. Das Gleiche gilt für die auditive und audiovisuelle Literaturrezeption, also vor allem die Arbeit mit Hörbuch und Film. Für die speziellen Belange des Erstlesens ist ferner auf DTP 2 „Schriftsprach- und Orthographieerwerb: Erstlesen, Erstschreiben“ hinzuweisen. Die vorliegenden drei lese- und literaturdidaktischen Bände gehen zwar auf alle Schulstufen von der Vorschule bis zur Sekundarstufe II ein; ausführliche Information zu den Methoden des Schriftspracherwerbs werden jedoch in DTP 2 gegeben.

Für die vorliegende zweite Auflage sind die Beiträge überarbeitet worden – mit Ausnahme des Beitrags von Karlheinz Fingerhut; dieser Beitrag ist unverändert, da Fingerhut verstorben ist und unseres Erachtens für den Beitrag auch kein Überarbeitungsbedarf bestand. An weiteren Veränderungen sind zu nennen: Der bisherige Beitrag B5.3 „PISA und Output-Orientierung: Standards und Evaluation“ ist ge-

strichen – die darin angesprochenen Fragestellungen sind, soweit noch aktuell, in den erweiterten Beitrag B5.1 aufgenommen, für den nach dem Tod von Karl Holle nun allein Steffen Gailberger verantwortlich zeichnet. Den Beitrag C4 „Kompetenzen und Unterrichtsziele im Lese- und Literaturunterricht der Sekundarstufe II“ hat Michael Kämper-van den Boogaart neu verfasst, den Beitrag E1 „Gegenwärtiger Stand der empirischen Unterrichtsforschung zur Vermittlung von Lesekompetenz“ Wolfgang Lenhard und den Beitrag F5 „Bewerten mündlicher Leistungen im Literaturunterricht“ Kerstin Hillegeist.

Berlin und Augsburg, 2009/2018

Michael Kämper-van den Boogaart
Kaspar H. Spinner

F

**Erfolgskontrollen
und Leistungsmessung**

JULIANE KÖSTER

F1 Aufgabentypen für Erfolgskontrollen und Leistungsmessung im Literaturunterricht

1 Was soll kontrolliert und gemessen werden?

Wer den Erfolg des Lese- und Literaturunterrichts kontrollieren und die Leistungen der Lerner messen will, sollte wissen, was kontrolliert und gemessen werden soll. Woran bemisst sich der Erfolg des Lese- und Literaturunterrichts? Welche Wissensbestände, welche Fähigkeiten, welche Fertigkeiten sollen ausgewiesen werden?

Geht es um die kognitiven Leistungsdispositionen, die den Leseprozess im strikten Sinn ausmachen, wie Rosebrock und Nix (2008) ihn in Abgrenzung zur Subjekt- und zur sozialen Ebene beschrieben haben, oder werden Fähigkeiten erwartet, wie Kaspar H. Spinner (2006) sie fürs literarische Verstehen formuliert? Während es auf der Prozessebene um ein Leistungsspektrum geht, das von der Wort- und Satzerkennung über die Etablierung globaler Kohärenz zur Identifikation von Superstrukturen und der Integration sprachlich-formaler Merkmale in das Gesamtverständnis reicht, beziehen die von Spinner formulierten Erwartungen sowohl die Ebene des Subjekts mit den emotionalen und motivationalen Komponenten ein als auch die soziale Ebene und den damit verbundenen Austausch zwischen Individuen. Die Fähigkeiten, „beim Lesen und Hören Vorstellungen [zu] entwickeln“ oder „subjektive Involviertheit und genaue Wahrnehmung miteinander ins Spiel [zu] bringen“, setzen die Subjektebene dominant. Entsprechendes gilt für die Erwartung, „sich auf die Unabschließbarkeit des Sinnbildungsprozesses ein[zulassen]“, „mit Fiktionalität bewusst um[zugehen“, „literaturhistorisches Bewusstsein [zu] entwickeln“ (Spinner 2006, 16 ff.). Die Bereiche der „Vorstellungen“ und des sich entwickelnden oder entwickelten „Bewusstseins“ sind über Leistungsaufgaben kaum zu erfassen. Anforderungen, wie z. B. „mit dem literarischen Gespräch vertraut [zu] werden“, schließen auch die soziale Ebene ein.

Anders verhält es sich mit der Fähigkeit, einen literarischen Text zu klassifizieren oder ihn historisch zu verorten. Anforderungen, die darauf zielen, sind überprüfungstechnisch gut zu bewältigen. Spinner spricht hier von „prototypische[n] Vorstellungen von Gattungen/ Genres“, die die Lerner gewinnen sollen. Wer solche Anforderungen erfüllt, muss aber einen literarischen Text als Ganzen noch nicht verstanden haben.

Wenn das Ziel im Kompetenzbereich „Lesen – Umgang mit Texten und Medien“ darin besteht, dass Texte als ganze verstanden werden, dann kommt alles darauf an, die Sicherung des inhaltlichen (Gesamt-)Zusammenhangs zu überprüfen. Die zentrale Frage lautet: Können Prüflinge darüber Auskunft geben, worum es in einem Text geht, oder beschränken sie sich auf die Angabe dessen, was darin vorkommt? Aus didaktischer Perspektive wird die große Bedeutung der Fähigkeit zur globalen Kohärenzbildung auch aus dem Umfeld der DESI-Studie bestätigt (vgl. Gailberger/Krelle/Triebl 2007).

Folgt man den Setzungen, wie sie in den BSMSA (Bildungsstandards für den Mittleren Schulabschluss) für das Fach Deutsch formuliert sind, dann gilt es, sowohl „Texte und Medien [zu] verstehen und [zu] nutzen“ als auch „Kenntnisse über Literatur [zu] erwerben“ (ebd., 8). Bezogen auf die erwarteten **Kenntnisse** finden sich dort folgende weitere Aussagen:

- „die eigenen Kenntnisse mit dem Thema, dem Hauptgedanken, der Problemstellung verbinden“ (ebd., 18),
- „inhaltliche und methodische Kenntnisse auf unbekannte Sachverhalte sinnvoll beziehen“ (ebd.).

Was das **Verstehen** angeht, so bieten die Bildungsstandards keinerlei Definition oder Erklärung. Michael Kämper-van den Boogaart hat wiederholt darauf hingewiesen, dass die einzelnen Standards für die Abschlüsse der Sekundarstufe I im Bereich „Literarische Texte verstehen und nutzen“ wenig mit dem Erwerb literarischer Verstehenskompetenz zu tun haben (Kämper-van den Boogaart 2005; 2008).

Lesekompetenztests wie PISA, PIRLS und DESI prüfen das Textverstehen auf der Prozessebene. Kompetenzen, die auch die Ebenen des Subjekts und der Anschlusskommunikation berühren, entziehen sich weitgehend der Messung. Mit anderen Worten: Leistungen, die in Abhängigkeit von Lesemotivation, Lesefreude, Interesse erbracht werden, also die Subjektebene einbeziehen, werden über die genannten Modelle nicht verlässlich erfasst.

Wenn es um den Erfolg des **Literaturunterrichts** geht, so stellt sich jedoch die Frage, was das Besondere am Verstehen literarischer Texte ist. Aus Spinners Katalog wären hier folgende Fähigkeiten zu nennen:

- Perspektiven literarischer Figuren nachvollziehen;
- Narrative und dramaturgische Handlungslogik verstehen;
- Metaphorische und symbolische Ausdrucksweise verstehen.

Auch hier fällt auf, dass der Verstehensbegriff nicht näher expliziert wird. Unklar bleibt, wodurch sich Verstehen und „Nachvollzug“ unterscheiden. Gleichwohl sind die von Spinner genannten Verstehensziele, nämlich die „Perspektiven der literarischen Figuren“, die „narrative und dramaturgische Handlungslogik“ sowie

„metaphorische und symbolische Ausdrucksweise“ zentrale Größen im schulischen Literaturunterricht.

Thomas Zabka sieht die Merkmale literarischer Rezeptionskompetenz im „Verstehen konkurrierender Information“ und ganz besonders in der „Erwartung und kognitive[n] Bewältigung systematischer Mehrdeutigkeit“ (Zabka 2006, 82 f.). Von systematischer Mehrdeutigkeit literarischer Texte ist zu sprechen, wenn diese „durch die Konkurrenz von Aussagen, Aussagenverknüpfungen und provozierten Inferenzen eine Pluralität der Verstehensmöglichkeiten eröffnen“ (Zabka 2006, 83).

Darüber hinaus hebt er neben „Informationsverknüpfung“ und „Inferenzbildung“, ohne die auch bei nichtliterarischen Texten keine globale Kohärenzbildung erfolgen kann, den Aspekt des „Verstehens von Indirektheit“ (ebd.) hervor.

Die von Zabka genannten Merkmale gehen von den grundlegenden kognitiven Prozessen der Informationsermittlung aus und dringen zu anspruchsvolleren Leistungen vor, die schließlich zur Deutung literarischer Texte führen. Der erste Punkt – „Informationsverknüpfung“ – steht ganz konkret im Dienste der Sicherung des inhaltlichen Zusammenhangs. „Inferenzbildung“ verdeutlicht, dass auf jeder Ebene des Verstehensprozesses Schlussfolgerungen gezogen werden müssen. „Erwartung und kognitive Bewältigung systematischer Mehrdeutigkeit“ und „Verstehen von Indirektheit“ sind schließlich für literarische Texte besonders relevant. Sie stellen einen hohen kognitiven Anspruch dar und sind in der Regel mit komplexen Reflexionsleistungen verbunden. Dabei ist die Sicherung des inhaltlichen Zusammenhangs bzw. die „Etablierung globaler Kohärenz“ immer die unverzichtbare Grundlage. Liest man Abiturarbeiten der Jahre 2007 bis 2017 im Hinblick auf die literarische Verstehenskompetenz der Prüflinge, so fällt auf, dass diese Anforderung von einem Teil der Prüflinge nicht oder nur partiell erfüllt wird. Zwar gelingt den meisten die Ermittlung und fachgerechte Benennung einzelner „sprachlich-stilistische[r] Merkmale eines Textes“, aber ungeachtet dieser punktuellen Leistungen hat eine Vielzahl der Prüflinge Probleme, „Sinnzusammenhänge zwischen einzelnen Einheiten dieser Texte her[zu]stellen und sie als Geflechte innerer Bezüge und Abhängigkeiten [zu] erfassen“ (BSAHR, 18).

Daraus wäre auch zu folgern, dass der damit verbundene kognitive Anspruch von Lehrpersonen häufig unterschätzt wird. Obwohl sich sowohl in den BSMSA als auch in den Bildungsstandards für die Allgemeine Hochschulreife (BSAHR) inhaltsbezogene Anforderungen finden, sind diese weniger ausgeprägt als die Anforderungen im Bereich der Diktion. Bezogen auf literarische Texte formulieren die BSMSA den Standard „zentrale Inhalte erschließen“ (BSMSA, 14) und die BSAHR verlangen in diesem Kompetenzbereich „Inhalt, Aufbau und sprachliche Gestaltung literarischer Texte [zu] analysieren, Sinnzusammenhänge zwischen einzelnen Einheiten dieser Texte her[zu]stellen und sie als Geflechte innerer Bezüge und Abhängigkeiten [zu] erfassen“ (BSAHR, 18).

Im Hinblick auf die Frage nach dem erfolgreich zu bewältigenden Gegenstand der Kontrollen und Messungen wäre festzuhalten, dass es sich um Texte handelt, die entweder literarische sind oder sich auf solche beziehen. Sofern die Kontrollen und Messungen auf Objektivität gerichtet sind, also vergleichbare Resultate intendieren, sollten sie sich einerseits auf die kognitiven Dispositionen beschränken, andererseits aber die spezifischen Facetten des literarischen Gegenstands berücksichtigen.

2 Leistungsaufgaben

Unter Aufgaben sind situationsbezogen formulierte Anforderungen zu verstehen. Die Lernsituation ist durch Erwerbs- und Übungsaufgaben bestimmt, die Leistungssituation durch Überprüfungsaufgaben. Deshalb ist es unbedingt erforderlich, zwischen Lern- und Leistungsaufgaben zu unterscheiden. Denn der **Erwerb** von Wissen und Können folgt anderen Gesetzen als deren **Überprüfung**. Diese Unterscheidung von Aufgaben ist weit verbreitet seit Weinert (1999) als Ergebnis von ihm betreuter Längsschnittstudien 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. So können Fehler in der Lernsituation fruchtbar sein und als Erkenntnismittel dienen. Erfolgreicher Unterricht braucht nach Weinert beides, und in der Wahrnehmung der Schülerinnen und Schüler möglichst getrennt: viele entspannte Gelegenheiten zum intensiven Lernen und genügend anspruchsvolle Leistungssituationen. Folglich erscheint es sinnvoll, für beide Situationen unterschiedliche Aufgaben zu modellieren. Lernaufgaben können sehr unterschiedlich sein, denn sie stehen in enger Beziehung zum jeweiligen Zieltyp des Lernens (vgl. Oser/Sarasin 2005, 11).

Leistungsaufgaben sollen feststellen, welcher Leistungsstand von Schülerinnen und Schülern erreicht wurde und welcher Förderbedarf folglich besteht. „Um hinreichende Genauigkeit (Messfehler-Reduzierung) und Aussagekraft (Validität) bei Leistungsaufgaben zu erreichen, müssen prinzipiell ähnliche Anforderungen wie an Testaufgaben erfüllt werden – wenn auch reduziert.“ (Reisse 2008, 319 f.)

Gemeinsam ist allen Leistungsaufgaben im Bereich des Lese- und Literaturunterrichts, dass sie mündliche oder schriftliche Texte zum Gegenstand haben und auf eine bestimmte Textverstehensleistung ausgerichtet sind. Konsens besteht wohl darin, zwischen drei Arten von Verstehensleistungen zu unterscheiden:

(1) Leistungen, die sich auf den Leseprozess beziehen. Das sind Leistungen im Bereich der Wort- und Satzerkennung und der Etablierung lokaler Kohärenz einerseits sowie Leistungen im Bereich der Etablierung globaler Kohärenz und der Identifikation bzw. Integration von Superstrukturen und Darstellungsstrategien andererseits (vgl. Rosebrock/Nix 2008). Diese prozessbezogenen Leistungen sind Bestandteil der Lesekompetenztests (PISA, DESI) und gehen als Analyse- und Interpretationsleistungen in Klassenarbeiten und Klausuren ein.

(2) Leistungen, die auf der Herstellung extratextueller Bezüge beruhen. Dazu gehört auch die explizite Verknüpfung des Ausgangstextes mit Vergleichstexten oder bestimmten Wissensbeständen. Eine solche Verknüpfung kann sowohl erläuternden als auch wertenden Charakter haben. PISA hat diese Fähigkeit im Rahmen von Aufgaben zum „Reflektieren und Bewerten“ erhoben.

(3) Leistungen, die die motivationale und affektive Beteiligung des Lesers einbeziehen. Individuelle Faktoren wie z. B. die Fähigkeit, „lebendige Vorstellungen beim Lesen und Hören literarischer Texte [zu] entwickeln“ (Bildungsstandards für den Primarbereich), lassen sich nur schwer oder gar nicht messen. Das gilt auch für „die Motivation, sich einem bestimmten Lesestoff zuzuwenden“ und „die Bereitschaft, sich vor allem bei literarischen Texten auch affektiv zu beteiligen“ (Köster/Rosebrock 2009, 115 f.). Obwohl diese Faktoren beim Verstehen literarischer Texte eine wichtige Rolle spielen, gibt es bislang keine Verfahren, sie den Testgütekriterien entsprechend zu erheben. Natürlich kann in einer Klassenarbeit Gelegenheit geboten werden, Vorstellungen, die anhand eines Textes gebildet wurden, zu formulieren oder über die affektive Beteiligung beim Lesen Auskunft zu geben. Die jeweils erbrachten Leistungen zu vergleichen und objektiv auszuwerten, dürfte jedoch auch in Klassenarbeiten und Klausuren Probleme bereiten.

3 Kategorien zur Beschreibung von Leistungsaufgaben

Unter Aufgabenarten, Aufgabentypen, Aufgabenformaten wird die „Art und Weise der Aufgabenstellung und der daraus folgenden Aufgabenbeantwortung“ (Reisse 2008, 318) verstanden. „Aufgabenformate beschreiben typische Kategorien von Aufgaben, unterschieden

- nach der Art der Beantwortung = antwortbezogene Aufgabenformate (z. B. gebundene oder offene schriftliche Aufgabenarten)
- oder nach den Kompetenzen, die mit ihnen erfasst werden sollen = kompetenzbezogene Aufgabenformate.“ (Ebd.)

3.1 Art der Beantwortung

In der Kategorie „Art der Beantwortung“ spielen folgende Unterscheidungen eine Rolle:

- **Darstellungsmodus/materiale Ausführung**

Besteht sie in einer Rezeptionsleistung ohne Textproduktion oder in einer Rezeptionsleistung mit Textproduktionsleistung?

- **Offenheit** (vs. Geschlossenheit) der Aufgabe (Entscheidungsspielraum)

Textproduktionsaufgaben sind grundsätzlich offene Aufgaben, auch wenn sie den Schülerinnen und Schülern unterschiedlich großen Entscheidungsspielraum gewähren. Der Auftrag, eine Erzählung zusammenzufassen, bedeutet eine stärkere Festlegung als der Auftrag, eine Erzählung fortzusetzen. Der allgemeine Interpretationsauftrag – „Interpretiere die vorliegende Szene aus ‘Emilia Galotti’“ – lässt den Schülerinnen und Schülern mehr Freiheit als die Spezifizierung dieses Auftrags – „Interpretiere die Szene im Hinblick auf die darin wirksamen Kommunikationsstrategien“.

3.2 Kompetenzen, die erfasst werden sollen

In der Kategorie „Kompetenzen, die mit ihnen erfasst werden sollen“ spielen folgende Unterscheidungen eine Rolle:

- hinsichtlich der **Komplexität** (vs. Linearität) der Aufgabe (Integrationsgrad)

Wenn die Überprüfung der Textverstehenskompetenz an eine umfängliche Schreibleistung gebunden wird, dann ist die Aufgabe in der Regel komplexer, als wenn dies durch Multiple-Choice-Aufgaben, Richtig-Falsch-Aufgaben oder Kurzantwort-Aufgaben geschieht. Die Komplexität einer Aufgabe bemisst sich an der Zahl der zu verknüpfenden Variablen und deren internem Verknüpfungsgrad (vgl. Funke 2006; Jonassen 2000). Folglich haben Aufgaben, deren Lösung die adaptive Verknüpfung mit Vorwissensbeständen oder Vergleichstexten erfordert, einen hohen Komplexitätsgrad und weisen höhere empirische Schwierigkeit auf.

- hinsichtlich der zur Aufgabenlösung erforderlichen **Schlussfolgerungsart** (Induktion vs. Deduktion)

Nachweisaufgaben auf der Basis einer allgemeinen Setzung sind der Deduktion verpflichtet, Ermittlungsaufgaben der Induktion. Sie verlangen vom Besonderen auf das Allgemeine zu schließen. „Weise nach/Zeige, dass es sich beim vorliegenden Text um ein Märchen handelt“ ist ein typisches Beispiel für Deduktion. Sie setzt voraus, dass den Prüflingen die Merkmale des Märchens bekannt sind und als Suchinstrument genutzt werden können. Demgegenüber besteht der exemplarische Ermittlungsauftrag darin, aus dem Vergleich mindestens zweier Märchen deren Gemeinsamkeiten zu erheben und daraus auf Märchenmerkmale zu schließen. Ein didaktisch durchaus relevantes „Zwischenglied“ bilden sog. Überprüfungsaufgaben: „Prüfe, um welche Textsorte es sich im vorliegenden Text handelt, und begründe deine Antwort mit konkreten Textstellen.“ Folgt man der von Iris Winkler vorgenommenen Unterscheidung¹ von **Rekonstruktionsaufgaben**, **Ge-**

¹ Diese Unterscheidung gilt sowohl für Lern- als auch für Leistungsaufgaben.

nerierungsaufgaben und **Diskussionsaufgaben**, dann gehört die Nachweisaufgabe zu den **Rekonstruktionsaufgaben**. Sie geben das Verstehensprodukt vor „und fordern vom Aufgabenbearbeiter die Rekonstruktion bzw. den Nachvollzug des gegebenen Textverstehensproduktes“ (Winkler 2011, 111). Ermittlungsaufgaben im o. g. Sinn sind **Generierungsaufgaben**. Sie verlangen vom Aufgabenbearbeiter die Bildung des Verstehensproduktes und geben als allgemeine Kategorie die Ebene des Textverstehens vor. Die vorgestellte Überprüfungsaufgabe zählt dann zu den **Bewertungsaufgaben**, „die ein Verstehensprodukt vor[geben] und [...] vom Aufgabenbearbeiter dessen Prüfung bzw. Bewertung anhand eines bestimmten Reflexionskriteriums“ fordern (ebd.).

4 Einsatz von Leistungsaufgaben

Vor diesem Hintergrund lassen sich **drei Gruppen** von Leistungsaufgaben hinsichtlich ihres Einsatzes unterscheiden:

1. Aufgaben für Klassenarbeiten und Klausuren zur Überprüfung des Unterrichtserfolges,
2. Aufgaben für Zentrale Prüfungen zur Zertifizierung von Abschlüssen und
3. normierte Testaufgaben für „Assessments zur Bilanzierung von Lernergebnissen“.

4.1 Aufgaben für Klassenarbeiten zur Überprüfung des Unterrichtserfolges

Aufgaben für Klassenarbeiten zu entwickeln, ist in erster Linie Sache der einzelnen Lehrperson. Sie ist frei, den Voraussetzungen der Schülerinnen und Schüler entsprechend Art und Typus der Aufgabe zu bestimmen. Hier spielen natürlich die länderspezifischen Traditionen ebenso eine Rolle wie Absprachen in den einzelnen Fachkonferenzen.

Bei Klassenarbeiten ist das Spektrum von Aufgabentypen sehr groß und variantenreich. Letztlich wissen wir wenig über das Spektrum und den Variantenreichtum von Aufgabentypen für Klassenarbeiten. Wir wissen auch wenig darüber, ob Klassenarbeiten überwiegend auf die Reproduktion von im Unterricht erworbenem literarischem Wissen zielen (z. B. „Nenne die Merkmale einer Ballade“), ob sie die Identifizierung der Balladenmerkmale an einer unbekannten Ballade verlangen (z. B. „Weise nach, dass Heines Gedicht ‘Belsazar’ eine Ballade ist“), ob sie die adaptive Nutzung des im Unterricht erworbenen Wissens und Könnens verlangt („Heine hat die Geschichte von Belsazar aus der Bibel übernommen. Vergleiche den vorliegenden Bibeltext mit der Ballade, ermittle, was Heine weggelassen oder verändert hat und stelle deine Ergebnisse in einer Tabelle dar“), ob sie Rezeptionsleistung und

Produktionsleistung getrennt erheben. Wenn zur Erhebung der Textverstehensleistung geschlossene Aufgaben eingesetzt werden, dann stellt sich die Frage, ob es sich in erster Linie um Lokalisierungsaufgaben handelt oder ob auch die Fähigkeit zur globalen Kohärenzbildung, zur Sinnerzeugung und zur Bewertung des Gelesenen gefordert wird.

Als Beispiel für typische Textverstehensaufgaben in Klassenarbeiten werden hier – getrennt nach Verstehens Ebenen (vgl. PISA 2000; Rosebrock/Nix 2006) – Aufgabenbeispiele vorgestellt, die den Lernerfolg am Ende einer Unterrichtsreihe über Balladen und Erzählgedichte überprüfen. Sie unterscheiden sich hinsichtlich des Abstraktionsniveaus, des Entscheidungsspielraums und des Komplexitätsgrads. Dabei werden sowohl Aufgaben berücksichtigt, die stärker an den unterrichtlichen Traditionen orientiert sind und sich weitgehend auf die Angabe eines Operators beschränken, als auch Aufgaben, die konkretere inhaltsbezogene Hilfen geben und auch geschlossene Formate berücksichtigen.

Bertolt Brecht

Die Krücken

Sieben Jahre wollt kein Schritt mir glücken.

Als ich zu dem großen Arzte kam

Fragte er: Wozu die Krücken?

Und ich sagte: Ich bin lahm.

Sagte er: Das ist kein Wunder.

Sei so freundlich, zu probieren!

Was dich lähmt, ist dieser Plunder.

Geh, fall, krieche auf allen vieren!

Lachend wie ein Ungeheuer

Nahm er mir die schönen Krücken

Brach sie durch auf meinem Rücken

Warf sie lachend in das Feuer.

Nun, ich bin kuriert: ich gehe.

Mich kurierte ein Gelächter.

Nur zuweilen, wenn ich Hölzer sehe

Gehe ich für Stunden etwas schlechter.

Verstehensebene	Aufgaben															
1. Lokalisieren	1a) Welche Handlungen lassen sich wem zuordnen? <table><tr><td></td><td>Arzt</td><td>Patient</td></tr><tr><td>Wer</td><td></td><td></td></tr><tr><td>lacht</td><td>☐</td><td>☐</td></tr><tr><td>nennt die Ursache des Leidens</td><td>☐</td><td>☐</td></tr><tr><td>hat das letzte Wort</td><td>☐</td><td>☐</td></tr></table> 1b) Schreibe auf, um welches Gebrechen es im Gedicht geht.		Arzt	Patient	Wer			lacht	☐	☐	nennt die Ursache des Leidens	☐	☐	hat das letzte Wort	☐	☐
	Arzt	Patient														
Wer																
lacht	☐	☐														
nennt die Ursache des Leidens	☐	☐														
hat das letzte Wort	☐	☐														
2. Globale Kohärenz bilden	2a) Gib den Inhalt des Gedichts mit eigenen Worten wieder. 2b) Ermittle im Text, wie der Arzt mit dem Gebrechen des Patienten umgeht und zu welchem Erfolg seine Behandlung führt. Schreibe deine Ergebnisse auf.															
3. Sinn konstruieren	3a) Interpretiere das Gedicht. 3b) Der Arzt im Gedicht gründet seine Heilmethode auf ☐ Grobheit gegenüber dem Patienten. ☐ Vertrauen in die Kraft des Patienten. ☐ Gelächter über den Patienten. ☐ Vernichtung des Eigentums des Patienten. Kreuze die beste Antwort an und begründe deine Entscheidung textbezogen. 3c) Überprüfe die These, dass das Gedicht ein Plädoyer für Mündigkeit ist. 3d) Erkläre, warum der Patient Probleme beim Gehen hat, wenn er „Hölzer“ sieht?															
4. Bewerten	4a) Nimm zum Gedicht [kritisch] Stellung. 4b) Würdest du diesen Arzt weiterempfehlen? Begründe deine Antwort. 4c) Setze dich mit der These auseinander, dass es sich um ein religionskritisches Gedicht handelt.															

Diese zehn Aufgaben sind auf unterschiedlichen Ebenen des Verstehens angesiedelt. Aufgaben der ersten Verstehensebene stellen Anforderungen unterhalb der Ebene globaler Kohärenzbildung dar. Das heißt: Man kann die Aufgabe lösen, ohne den Text als Ganzen verstanden zu haben. Während Aufgaben auf dieser Verstehens-ebene in Lesetests eine prominente Rolle spielen, sind sie in Klassenarbeiten eher selten. Wenn relevante Informationen ermittelt werden sollen, dann helfen derartige Anforderungen vor allem schwächeren Lesern bei der Sicherung des inhaltlichen Zusammenhangs. Für literarische Texte gilt allerdings, dass lokale Informationen

häufig den Rekurs auf das Globale erfordern. Die Frage nach dem Gebrechen, von dem im Gedicht die Rede ist, lässt sich jedoch ohne Verständnis des Gesamtzusammenhangs ermitteln.

Die Aufgaben auf der dritten und vierten Verstehensebene setzen immer die Etablierung globaler Kohärenz voraus. Deshalb kommt diesem Prozess (auf der zweiten Verstehensebene) große Bedeutung zu. Die erste Aufgabe „Gib den Inhalt des Gedichts mit eigenen Worten wieder“ ist schwierig, weil die Anforderung ganz formal bleibt und weder einen Hinweis auf Wichtiges noch Strukturierungshilfen bietet. Die Schülerinnen und Schüler müssen allein über Auswahl und Anordnung der Informationen entscheiden. Anders verhält es sich mit der Aufgabe „Ermittle im Text, wie der Arzt mit dem Gebrechen des Patienten umgeht und zu welchem Erfolg seine Behandlung führt. Schreibe deine Ergebnisse auf.“ Hier wird ein zweischrittiges Vorgehen angeboten und die Aufmerksamkeit zunächst auf das Verhalten des Arztes und dann auf den Erfolg seiner Bemühungen gerichtet. Dadurch, dass die Aufgabe 2a den Ausweis der Kohärenzetaablierung mit eigener Textproduktion verbindet, ist sie deutlich schwerer zu bewältigen als die Aufgabe, das Verhalten des Arztes und dessen Erfolg in wenigen Sätzen zu sichern. Die Aufgaben auf der dritten Ebene zielen auf Sinnkonstruktion und Interpretation. Auch hier stellt der unspezifizierte Interpretationsauftrag die schwierigste Anforderung dar.

Auf der Ebene der Bewertung des Textes werden die Schülerinnen und Schüler in Aufgabe 4a zur eigenen Stellungnahme aufgefordert. Oft wird auch eine kritische Stellungnahme erwartet. Schülerinnen und Schülern bereitet dieser Auftrag häufig deshalb Schwierigkeiten, weil ihnen Kriterien bzw. begründete Wertmaßstäbe fehlen, um sich produktiv mit dem jeweiligen Sachverhalt auseinanderzusetzen. Der offene Auftrag „Stellung zu nehmen“ schließt nämlich ein, dass die Schülerinnen und Schülern selbst relevante Aspekte der Kritik finden. Häufig erweist sich das aber als Überforderung.

Die Aufgabe 4b setzt hier geringere Ausgangskompetenz voraus: „Würdest du diesen Arzt weiterempfehlen? Begründe deine Antwort.“ Ohne den **Entscheidungsspielraum** zu eliminieren, wird das **Untersuchungsfeld** durch die Vorgabe einer konkreten Entscheidungssituation stärker begrenzt und die zu verknüpfenden **Variablen** werden deutlich reduziert (vgl. Köster 2005). Entsprechendes gilt auch für Aufgabe 4c, deren größere Schwierigkeit darin besteht, dass sie Wissen über Religionskritik voraussetzt. Je näher jedoch zentrale Abschlussprüfungen rücken, desto stärker orientieren sich Klassenarbeiten an den von der KMK festgelegten Aufgabenarten für zentrale Prüfungen.

4.2 Aufgabenarten für zentrale Prüfungen

Für zentrale Prüfungen – den Hauptschulabschluss, den Mittleren Schulabschluss und das Abitur – werden Art und Typus der Aufgaben durch die Bildungsadministra-

tion (KMK und Kultusministerien der Länder) vorgegeben. Die damit verbundene Aufgabenkonzeption wird in den Bildungsstandards für den Mittleren Schulabschluss folgendermaßen erläutert:

„Charakteristisch für diese Aufgabenkonzeption ist ein Ansatz, der die Kompetenzbereiche ‘Umgang mit Texten und Medien’, ‘Schreiben’ sowie ‘Sprache und Sprachgebrauch’ in ihrer Komplexität und die einzelnen Standards in unterschiedlicher Intensität und Breite erfasst. Daher sind die konkreten Aufgabenstellungen vorherrschend auf die Bearbeitung eines Textes (oder Textensembles) bezogen und setzen die Fähigkeit voraus, über die grundlegenden/wesentlichen Aspekte des Textverstehens – vom allgemeinen Textverständnis über Informationsentnahme, Textinterpretation bis hin zur Reflexion und Bewertung von Inhalt und Form des Textes – zu verfügen. Jede der komplexen Aufgaben verlangt in der Bearbeitung, diesen Prozess zu durchlaufen. Reflexions- und Bewertungsleistung können auch in der Form eines eigenständig gestaltenden Umgangs oder einer argumentativen Auseinandersetzung mit dem Text oder einzelnen Textaussagen erfasst werden.“ (BSMSA, 20) Daraus resultieren fünf Aufgabenarten für zentrale Abschlussprüfungen. Sie gelten sowohl für den Mittleren Schulabschluss als auch für den Hauptschulabschluss.

4.2.1 Fünf Aufgabenarten

- einen [literarischen] Text untersuchen, analysieren, interpretieren,
- von einer Textgrundlage ausgehend [einen literarischen Text bzw. ein Textproblem] erörtern,
- eine Problemstellung erörtern [die aus einem literarischen Text resultiert]
- einen [literarischen] Text umformen oder weiterschreiben,
- von einer Textgrundlage ausgehend einen [literarischen] Text gestalten, entwerfen.

Die Standards für die Allgemeine Hochschulreife unterscheiden zwischen textbezogenem und materialgestütztem Schreiben. Im Bereich „Textbezogenes Schreiben“ sind vier Aufgabenarten festgelegt und definiert: Interpretation literarischer Texte, Analyse pragmatischer Texte, Erörterung literarischer Texte, Erörterung pragmatischer Texte. Im Bereich „Materialgestütztes Schreiben“ werden zwei Aufgabenarten unterschieden: Materialgestütztes Verfassen informierender Texte und Materialgestütztes Verfassen argumentierender Texte (BSAHR: 24–26). Anders als beim textbezogenen Schreiben, bei dem das Schreiben die Rezeptionsleistung dokumentiert, steht beim materialgestützten Schreiben die Rezeption im Dienst der Textproduktion. Die Länder haben sich darauf verständigt, dass die Prüfungsaufgaben im Abitur „einer der im Fach Deutsch möglichen Aufgabenarten in ihrer jeweiligen Definition“ entsprechen (IQB 2015). Bei Mischformen müsse die dominierende

Aufgabenart deutlich werden (vgl. ebd.). Seit 2017 steht für die Abiturprüfung ein ländergemeinsamer zentraler Aufgabenpool zur Verfügung und die Länder haben sich verpflichtet, Aufgaben zu entnehmen.

4.2.2 Drei Arbeitsanweisungen: Makrooperatoren

Für die Bearbeitung der einzelnen Aufgabenarten gelten bei den Abschlüssen der Sekundarstufe I die folgenden zentralen Arbeitsanweisungen:

- Untersuchen, Analysieren, Interpretieren,
- Informieren, Argumentieren, Erörtern,
- Gestalten, Entwerfen.

Bei zentralen Prüfungen sind nicht nur die Aufgaben, sondern auch die entsprechenden Lösungserwartungen und Korrekturanleitungen vorgegeben. Sie werden von – durch die Kultusministerien beauftragten – Deutschlehrkräften entwickelt. Bei komplexen offenen Interpretations- und Erörterungsaufträgen erweist sich die Entwicklung von Lösungserwartungen nicht zuletzt deshalb als problematisch, weil solche Aufgaben zu einer Lösungsvielfalt führen, die sich in einem „Erwartungshorizont“ kaum darstellen lässt. 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. Der Anspruch, den Korrektoren deutliche Orientierung zu geben, ohne unzulässige Festlegungen zu treffen, erscheint bei komplexen und offenen Aufgaben nur schwer einlösbar (Köster 2006; Kammler 2008).

4.2.3 Anforderungsbereiche und Operatoren

In allen Fächern unterscheiden die Bildungsstandards für alle Stufen und Abschlüsse drei Anforderungsbereiche. Bereits in den Erprobungsfassungen der „Einheitliche[n] Prüfungsanforderungen in der Abiturprüfung“ (1975 ff.) wurde nicht nur zwischen „inhaltsbezogenen Kenntnissen und Fähigkeiten“ und „methodenbezogenen Kenntnissen und Fähigkeiten“ unterschieden, sondern auch zwischen drei „Lernzielkontrollen“, die die Makro-Operatoren „Wiedergeben“, „Zusammenhänge herstellen“, „Reflektieren und beurteilen“² einführten und damit die Unterscheidung zentraler kognitiver Anforderungen ermöglichten. Diese Vorgaben wurden allerdings als „Normierungsversuch“ wahrgenommen und haben „heftige Diskussionen und kraftvolle Abwehrkämpfe“³ provoziert. Das führte zu einem Jahrzehnte währenden Stillstand. Erst in der Diskussion der PISA-Leseaufgaben wurden die unterschiedlichen Dimensionen kognitiver Tätigkeit wieder auf-

² http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_10_15-Bildungsstandards-Deutsch-Primar.pdf, S. 17

³ Dieter Lenzen (2012): Lehrerbildung in Deutschland – Lost in Translation. <https://www.uni-hamburg.de/uhh/organisation/praesidium/praesident/lehrerbildung-deutschland.pdf>, S. 3.

gegriffen und sind seit 2003 in allen Bildungsstandards verankert (vgl. Köster 2016, 48–55).

In den **Bildungsstandards für den MSA** heißt es entsprechend:

„Für Aussagen über die Angemessenheit, Qualität und Komplexität der Anforderungen, die mit den im Folgenden vorgestellten schriftlichen und mündlichen Aufgaben verbunden sind, stellen die Anforderungsbereiche einen Orientierungsrahmen dar, in dem sich die Leistungen von Schülerinnen und Schülern erfahrungsgemäß bewegen.“ (BSMSA, 17)

Anforderungsbereich I:

Verfügbarkeit der für die Bearbeitung der Aufgaben notwendigen inhaltlichen und methodischen Kenntnisse,

Anforderungsbereich II:

Selbständiges Erfassen, Einordnen, Strukturieren und Verarbeiten der aus der Thematik, dem Material und der Aufgabenstellung erwachsenden Fragen/Probleme und deren entsprechende gedankliche und sprachliche Bearbeitung,

Anforderungsbereich III:

Eigenständige Reflexion, Bewertung bzw. Beurteilung einer komplexen Problemstellung/Thematik oder entsprechenden Materials und ggf. die Entwicklung eigener Lösungsansätze. (Ebd., 10 f.)

In den BSAHR Deutsch:

„Anforderungsbereich I

umfasst das Wiedergeben von Sachverhalten und Kenntnissen im gelernten Zusammenhang, die Verständnissicherung sowie das Anwenden und Beschreiben geübter Arbeitstechniken und Verfahren.

Anforderungsbereich II

umfasst das selbstständige Auswählen, Anordnen, Verarbeiten, Erklären und Darstellen bekannter Sachverhalte unter vorgegebenen Gesichtspunkten in einem durch Übung bekannten Zusammenhang und das selbstständige Übertragen und Anwenden des Gelernten auf vergleichbare neue Zusammenhänge und Sachverhalte.

Anforderungsbereich III

umfasst das Verarbeiten komplexer Sachverhalte mit dem Ziel, zu selbstständigen Lösungen, Gestaltungen oder Deutungen, Folgerungen, Verallgemeinerungen, Begründungen und Wertungen zu gelangen. Dabei wählen die Schülerinnen und Schüler selbstständig geeignete Arbeitstechniken und Verfahren zur Bewältigung der Aufgabe, wenden sie auf eine neue Problemstellung an und reflektieren das eigene Vorgehen.“ (BSAHR, 22)

Ungeklärt ist, ob die Anforderungsbereiche, wie die KMK sie in den Bildungsstandards als Maßstab für die Konstruktion von Leistungsaufgaben vorgibt, **Schwierigkeitsgrade** anzeigen oder ob sie sich lediglich auf **unterschiedliche Ebenen des Textverstehens** beziehen. Wenn Ersteres der Fall ist, dann ist zu berücksichtigen, dass nicht jedes „selbstständige Urteilen und Bewerten“ eine höhere Anforderung darstellt als die Beschreibung der Argumentation eines Textes. Übereinstimmung besteht jedoch darin, dass Aufgaben für Zentrale Prüfungen alle drei Anforderungsbereiche berücksichtigen müssen.

Der Auftrag, einen literarischen Text zu interpretieren, schließt Leistungen in allen drei Anforderungsbereichen ein.

4.2.4 Probleme beim offenen Analyse- und Interpretationsauftrag

Will man die literarische Kompetenz von Schülerinnen und Schülern zu einem bestimmten Zeitpunkt ihrer Schulkarriere erheben, dann sind die Ergebnisse immer durch die eingesetzten Aufgaben gesteuert.

Bei offenen und zugleich komplexen Aufgabenstellungen in der Abiturprüfung zeigten sich schülerseitig häufig Probleme, die in erster Linie sowohl auf die fehlende Sicherung des gesamten inhaltlichen Zusammenhangs (Etablierung globaler Kohärenz) zurückzuführen sind als auch auf Probleme im Bereich des Vorwissens. Auch wenn die selbstregulierte Erschließung literarischer Texte ein übergreifendes Ziel von Literaturunterricht ist, so kann die Studie von Michael Steinmetz (2013) über das Können von Abiturienten im Fach Deutsch doch empirisch erhärten, dass beim offenen, unspezifischen Interpretationsauftrag die Passung zwischen Anforderung und Kompetenzen der Schülerinnen und Schüler häufig nicht gegeben ist. Fehlt die Sicherung des inhaltlichen Zusammenhangs, dann kommt es zu verfrühten Deutungsentscheidungen. Problematische Textdeutungen jedoch provozieren ihrerseits defizitäre Textwahrnehmung und verhindern damit eine textadäquate Inhalts-sicherung.

In Bezug auf die Nutzung von Wissensbeständen zieht der schematische Rückgriff auf Vorwissensbestände sowohl fehlerhafte Textwahrnehmung als auch Fehldeutungen nach sich. Ebenso führen fehlende Wissensbestände zu problematischen bzw. unangemessenen Textdeutungen. Die Schwierigkeit, Vorwissen flexibel zu nutzen, sollte bei der Entwicklung von Prüfungsaufgaben bedacht werden. Sinnvoll wäre eine Modifizierung des Aufgabentyps „einen [literarischen] Text untersuchen, analysieren, interpretieren“, z.B. in Form von Hinweisen auf das zu aktivierende Wissen.

Neben der Aktivierung von Vorwissen spielt die Sicherung des inhaltlichen Zusammenhangs sowohl lokal als auch global (Etablierung globaler Kohärenz) eine wichtige Rolle. Denn die Sicherung des inhaltlichen Zusammenhangs ist notwendige Grundlage für eine plausible Interpretation. Das besondere Augenmerk auf die Ver-

stehensleistung in einer Prüfungsarbeit und auf die (paraphrasierende) Darstellung inhaltlicher Gesichtspunkte ist von der Hypothese geleitet, dass viele Fehlleistungen in der Deutung sich bei näherem Hinsehen schon als (Folge einer) Fehlleistung im inhaltlichen Nachvollzug der Texte entpuppen.

4.2.5 Zwei unterschiedliche Beispiele für Prüfungsaufgaben

Beispiel 1: Saarland, Mittlerer Schulabschluss 2008

Interpretation zur Pflichtlektüre

Heinrich Böll: Die verlorene Ehre der Katharina Blum

KATHARINA BLUM:

„... und ich hab' die Pistole rausgenommen und sofort auf ihn geschossen. Zweimal, dreimal, viermal. Ich weiß nicht mehr genau.“ (S. 135)

- Stelle dar, wie es zu dieser Tat Katharinas kommen konnte.
- Bedenke bei deinen Ausführungen die Lebensumstände Katharinas und das Verhalten der Vertreter der ZEITUNG, des Staates, der Wirtschaft, der Kirche und die Reaktionen der Hausbewohner und Zeitungsleser.
- Beurteile im Schlussteil die Tat Katharinas.

Während der Prüfung kannst du das Rechtschreibwörterbuch benutzen. Denke daran, dass du für die Rechtschreibung bis zu 6 Punkte erhalten kannst.

Beispiel 2: Schleswig-Holstein Beispielaufgaben, Mittlerer Schulabschluss

A Aufgaben zum Text- und Sprachverständnis (Teilaufgaben 13–15)

Textgrundlage: Heinrich Heine: Der tugendhafte Hund (1. Strophe)

Quelle: http://za.schleswig-holstein.de/docs/2017/uebungsheft_msa_17/MSA_uebungsheft_2017.pdf

Lies den folgenden Text. *Heinrich Heine: Der tugendhafte Hund (1. Strophe)*⁴

*Ein Pudel, der mit gutem Fug⁵
Den schönen Namen Brutus⁶ trug,
War vielberühmt im ganzen Land
Ob seiner Tugend und seinem Verstand.
Er war ein Muster der Sittlichkeit,
Der Langmut und Bescheidenheit.
Man hörte ihn loben, man hörte ihn preisen
Als einen vierfüßigen Nathan den Weisen⁷.
Er war ein wahres Hundewer!*

⁴ Heinrich Heine, Der tugendhafte Hund, Reclam: Stuttgart 1995, S. 262–264.

⁵ mit gutem Fug = mit voller Berechtigung

⁶ römischer Politiker

⁷ Nathan = literarische Figur, die für Menschlichkeit steht

*So ehrlich und treu! Eine schöne Seel!
 Auch schenkte sein Herr in allen Stücken
 Ihm volles Vertrauen, er konnte ihn schicken
 Sogar zum Fleischer. Der edle Hund
 Trug dann einen Hängekorb im Mund,
 Worin der Metzger das schöngehackte
 Rindfleisch, Schafffleisch, auch Schweinefleisch packte. –
 Wie lieblich und lockend das Fett gerochen,
 Der Brutus berührte keinen Knochen,
 Und ruhig und sicher, mit stoischer⁸ Würde,
 Trug er nach Hause die kostbare Bürde⁹.*

A 13 Der Text lässt sich inhaltlich in zwei Teile gliedern (Vers 1 bis 10 und Vers 11 bis 20).

Erläutere, worin sie sich inhaltlich unterscheiden.

/2 P.

A 14 Die Sprache des Gedichts ist altertümlich und poetisch.

Ersetze die fett gedruckten Ausdrücke durch heute übliche, allgemein verständliche Formulierungen.

/3 P.

Aussagen	
(Der Pudel) war vielberühmt im ganzen Land ob seiner Tugend und seinem Verstand.	
Er war ein Muster der Sittlichkeit .	
Er war ein wahres Hundejuwel .	

A 15 Lies den folgenden Textausschnitt.

Auch schenkte sein Herr in allen Stücken ihm volles Vertrauen. Er konnte ihn schicken sogar zum Fleischer. Der edle Hund trug dann einen Hängekorb im Mund, worin der Metzger das schöngehackte Rindfleisch (...) packte.

Heine verwendet bei der Darstellung des Hundes das Prinzip der Steigerung

Unterstreiche in dem Textausschnitt das Wort, das die Steigerung anzeigt. /2 P

Während die Prüfungsaufgabe aus dem Saarland am Ansatz der Bildungsstandards orientiert ist und die Kompetenzbereiche „Umgang mit Texten und Medien“ und

⁸ gelassen

⁹ Last

„Schreiben“ in drei komplexen Teilaufgaben verbindet, werden in Schleswig-Holstein Textrezeption und Textproduktion getrennt geprüft. Die Aufträge im Saarland beziehen sich lediglich auf die Inhalte der Erzählung und sparen die Darstellungsebene aus. Demgegenüber spielt in Schleswig-Holstein auch die sprachliche Gestaltung der Textvorlage eine Rolle.

4.2.6 Diskussion

Eine deutschdidaktische Diskussion der von der Bildungsadministration vorgegebenen Aufgabentypen liegt nur für „Analyse/Interpretation“ einerseits und für „Erörterung“ andererseits vor (zusammengefasst von Zabka in diesem Band). Die didaktischen Debatten gelten in erster Linie den traditionellen schulischen Textsorten (Nacherzählung, Inhaltsangabe, Analyse, Interpretation, Charakteristik, Erörterung: vgl. Zabka in diesem Band) und lassen den Bezug auf die konkreten Realisierungsprobleme, die sich in den Prüfungsarbeiten zeigen, außer Acht.

Die konkreten Aufgaben für den Mittleren Schulabschluss, wie sie hier exemplarisch aus zwei Ländern vorgestellt werden, weisen allerdings darauf hin, dass die in den Bildungsstandards verbindlich vorgegebenen Aufgabenarten Anpassung an die realen Fähigkeiten der Prüflinge verlangen.

4.3 Normierte Testaufgaben

Im Bereich der **Testaufgaben** sind Aufgaben für Individualdiagnostische Tests von Aufgaben für „Assessments zur Bilanzierung von Lernergebnissen“ zu unterscheiden. Die Entwicklung von Aufgaben für Assessment-Programme sind psychometrisch geleitet (IQB, IQSH). Sie unterwerfen die von Deutschlehrkräften unter Einbeziehung von Fach-Expertinnen und -Experten aus den Hochschulen generierten domänenspezifischen Aufgaben (Lesen, Schreiben, Zuhören) einer strikten Pilotierungs- und Bewertungsprozedur. Durch die Pilotierung mit einer hinreichend großen Zahl von Stichprobenschulen werden die Aufgaben nach bestimmten Kriterien analysiert. „Wenn nur sehr wenige Schülerinnen und Schüler eine Aufgabe lösen konnten, ist sie nicht geeignet. Das Gleiche gilt, wenn alle Schülerinnen und Schüler eine Aufgabe gelöst haben – dann ergibt sich kein differenziertes Bild. Neben der Lösungshäufigkeit ist ein weiteres Auswahlkriterium die Information darüber, ob sie zuverlässige und aussagefähige Hinweise über die zu überprüfenden Kompetenzen geben.“ (<http://vera.lernnetz.de/content/faq7.php>)

Um die Testaufgabenentwicklung nach wissenschaftlichen Standards zu gewährleisten, hat das IBQ (Institut für Qualitätsentwicklung im Bildungswesen) „Guidelines zur Entwicklung und Codierung guter Aufgaben“ bereitgestellt. Es wäre sinnvoll, diese Guidelines für Lehrkräfte transparent zu machen und sie auch im deutschdidaktischen Kontext zu diskutieren. Vor allem aber wird die inhaltliche, d.h. gegenstandsbezogene Qualität dieser Aufgaben ebenso kritisch zu analysieren

sein wie die von den einzelnen Ländern verantworteten Aufgaben für die Zentralen Prüfungen. Denn es gibt deutliche Anzeichen, dass es schwierig ist, gute Items zu entwickeln, die auf den Ausweis globaler Kohärenzbildung und der Sinnerzeugung gerichtet sind.

4.3.1 Aufgabenformate in Testaufgaben

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 1970-er Jahren wurde diese Debatte v.a. in der Auseinandersetzung mit dem „Aufsatz“ als Prüfungsform von Ingenkamp (1971) angestoßen und sie stellt für die Deutschdidaktik vor allem deshalb eine Herausforderung dar, weil die Offenheit einer Aufgabe per se als Qualitätsmerkmal gilt.

Zu den geschlossenen Aufgabenformaten zählen: Richtig-/Falsch-Items, Multiple-Choice-Items, Zuordnungs-Items, Umordnungs-Items (vgl. auch Bremerich-Vos 2003).

Als Beispiele werden im Folgenden Items vorgestellt, die sich auf James Thurber, „Die ziemlich intelligente Fliege“ beziehen und am IQSH für VERA 6 (2008) entwickelt wurden.

Die ziemlich intelligente Fliege

Eine große Spinne hatte in einem alten Haus ein schönes Netz gewoben, um Fliegen zu fangen. Jedesmal, wenn eine Fliege sich auf dem Netz niederließ und darin hängenblieb, verzehrte die Spinne sie schleunigst, damit andere Fliegen, die vorbeikamen, denken sollten, das Netz sei ein sicherer und gemütlicher Platz. Eines Tages schwirrte eine ziemlich intelligente Fliege so lange um das Netz herum, ohne es zu berühren, daß die Spinne schließlich hervorkroch und sagte: „Komm, ruh dich ein bißchen bei mir aus.“ Aber die Fliege ließ sich nicht übertölpeln.

„Ich setze mich nur an Stellen, wo ich andere Fliegen sehe“, antwortete sie, „und ich sehe bei dir keine anderen Fliegen.“

Damit flog sie weiter, bis sie an eine Stelle kam, wo sehr viele Fliegen saßen. Sie wollte sich gerade zu ihnen gesellen, als eine Biene ihr zurief: „Halt, du Idiot, hier ist Fliegenleim. Alle diese Fliegen sitzen rettungslos fest.“

„Red keinen Unsinn“, sagte die Fliege. „Sie tanzen doch.“

Damit ließ sie sich nieder und blieb auf dem Fliegenleim kleben wie all die anderen Fliegen. [...]

(Thurber, James. 75 *Fabeln für Zeitgenossen*. Hamburg: Rowohlt 1967, 8)

Richtig-/Falsch-Items:

Kreuze an, ob die folgenden Aussagen richtig oder falsch sind.

	Richtig	Falsch
A Die Spinne ist hinterhältig.		
B Die Spinne hat auf ihrem Netz Fliegenleim aufgetragen.		

Multiple-Choice-Items:**Welche andere Überschrift passt zur Fabel?****Kreuze die richtige Antwort an.**

- A: Die Biene, die gemein und listig ist
 B: Die Fliege, die glaubt, clever zu sein
 C: Die Spinne, die alle Fliegen überlistet
 D: Der Fliegenleim, der Glück bringt

Zuordnungs-Items:**Ordne den Tieren die unten stehenden Handlungen zu.****Handlungen:**
 verzweifelter Befreiungsversuch – Nahrungsfang – Ratschläge ablehnen –
 Warnung

Tier	Handlung
A die Spinne	
B die ziemlich intelligente Fliege	
C die Biene	
D die anderen Fliegen	

Umordnungs-Items:**Ordne die folgenden Handlungsabschnitte so an, wie sie im Text erscheinen.****Schreibe die Nummern von 1 bis 7 in die jeweils passende Zeile.**

A Die Spinne will die Fliege überreden, sich bei ihr auszuruhen.	
B Die Spinne webt ein Netz.	
C Die Biene warnt die Fliege.	
D Die Fliege sieht das Spinnennetz als Gefahr an.	
E Die Fliege schlägt die Warnung aus.	
F Die Fliege klebt fest.	
G Die Spinne frisst die gefangenen Fliegen.	

Als halboffen gelten Kurzantwort-Items und Lückentexte. Z. B.

Der Text handelt von drei , die sich wie verhalten.

Überall da, wo die Lerner kleinere oder größere Texte mündlich oder schriftlich produzieren, handelt es sich um offene Formate. Z. B.

Versetze dich in die Rolle der Biene. Schreibe in einem Satz auf, was du als Biene zur Fliege sagen würdest, als diese am Fliegenleim kleben bleibt.

Wovon hängt die Option für ein bestimmtes Format ab? Geschlossene und halboffene Formate haben den Vorteil hoher Auswertungsobjektivität. Das heißt, dass die Übereinstimmung der Kodierer bzw. Korrektoren sehr hoch ist und der zeitliche Aufwand gering. Bei geschlossenen Items kann die Korrektur auch maschinell bzw. rechnergestützt erfolgen. Aber nicht alle Verstehensebenen lassen sich gleich gut über geschlossene bzw. halboffene Formate überprüfen. Lesetests mit wenig oder gar keinen offenen Items bieten oft nur wenig Gelegenheit, die Etablierung globaler Kohärenz oder Deutungsleistungen zu überprüfen. Folglich bieten sich offene Formate vor allem unter zwei Bedingungen an:

(1) Wenn es gilt, **bestimmte Teilkompetenzen** auszuweisen, nämlich die Fähigkeit zur Sinnerzeugung bzw. zur Deutung literarischer Texte.

Zum einen geben offene Aufgaben dem Bearbeiter mehr Entscheidungsspielraum als halboffene oder geschlossene Items. Zum anderen ist es überaus schwierig, MC- oder Richtig-Falsch-Items zu generieren, die auf Deutungsleistungen zielen. Denn die unzutreffenden Angebote müssen einerseits plausibel, andererseits aber falsch sein. Da Deutungsleistungen aber zu großen Teilen Sache der Plausibilität sind, hat man ein Problem, falsche Angebote hervorzubringen.

(2) Wenn es gilt, einen **hohen Ausprägungsgrad** einer bestimmten Kompetenz auszuweisen.

Folgt man den am IQB erzielten Ergebnissen, so wird die Fähigkeit, Informationen selbst zu erzeugen, z. B. als Zusammenfassung eines Textes oder als Formulierung des Textthemas, auf einen hohen Ausprägungsgrad der (Teil-)Kompetenz zurückgeführt. Entsprechend wurde dort für Schülerinnen und Schüler des 9. und 10. Jahrgangs festgestellt, dass zahlreiche Aufgaben auf Anforderungsniveau 5 verlangen, Textinformation „in längeren, inhaltlich und/oder sprachlich komplexen Texten wieder[zug]eben und [zu] verarbeiten“ (iKSM_Lesen_2016_09_08.pdf, 15).

Dabei ist allerdings zu bedenken, dass Rezeptionsaufgaben, die mit einer Textproduktionsleistung verbunden sind, eine hybride Kompetenz überprüfen, nämlich zu verstehen und das Verstandene schriftlich zu verbalisieren. Dabei hängt die gemessene Rezeptionsleistung von der Produktionskompetenz ab. Aus psychometrischer Sicht stellt das ein Problem dar.

Zudem bestehen offene Aufgaben häufig die Pilotierung nicht. Wenn die Übereinstimmung der Korrektoren bzw. Codierer weniger als 80 Prozent beträgt, sollte das Item nicht eingesetzt werden.

4.3.2 Gütekriterien

Welche Gütekriterien müssen Testaufgaben erfüllen?

Anders als die Aufgaben für Klassenarbeiten und Zentrale Abschlussprüfungen werden die in den Vergleichsarbeiten verwendeten Aufgaben nach den allgemein geltenden Kriterien der Testgüte – Objektivität, Reliabilität, Validität, Fairness und angemessene Schwierigkeit – erarbeitet und nach einer Erprobung ausgewählt:

- Unter **Objektivität** wird verstanden, dass die Bewertung durch unterschiedliche Personen für dieselbe Schülerantwort gleich ausfällt. Um dies zu gewährleisten, werden in den Vergleichsarbeiten vor allem Multiple-Choice-Aufgaben, Richtig/Falsch-Aufgaben und Zuordnungsaufgaben eingesetzt. Aufgaben, die freie Antworten der Schülerinnen und Schüler erfordern, benötigen eine ausführliche Auswertungsanleitung, um eine hinreichende Übereinstimmung der Aufgabenbeurteilung zu erzielen.
- Die **Reliabilität** bezeichnet die Genauigkeit des Testergebnisses. Gestützt auf die Erprobungsdaten werden nur Aufgaben, die einem Mindeststandard an Diagnosegenauigkeit genügen, für die Vergleichsarbeit zugelassen.
- Dass der Test tatsächlich den zu prüfenden Testbereich erfasst, wird mit der **Validität** festgestellt. In den Vergleichsarbeiten werden dazu alle Aufgaben den einzelnen Komponenten der Bildungsstandards zugeordnet und ihre Einordnung von Expertinnen und Experten geprüft.
- Eine hinreichende **Fairness** besteht dann, wenn die Aufgabenzusammenstellung die Interessen von Jungen und Mädchen gleichermaßen berücksichtigt und keine speziellen kulturellen Inhalte voraussetzt, die Jugendliche anderer Herkunft benachteiligen.

Die Aufgabendiskussion der 1970-er Jahre hatte vor allem das Kriterium der **Auswertungsobjektivität** (d.h. eine bestimmte Schülerlösung muss immer gleich bewertet werden, unabhängig davon, wer die Aufgabe auswertet) im Blick, die bei offenen Aufgaben nach wie vor eine beträchtliche Herausforderung (genaue und detaillierte Hinweise, was als richtig und was als falsch zu werten ist) darstellt und ihrerseits mit zunehmender Komplexität einer Aufgabe wiederum wächst. Angesichts der Vielzahl benötigter Items ist immer auch die **Validität** zu reflektieren: Prüft die Aufgabe, was sie prüfen soll? So prüfen offene Items nicht nur, ob der Leser die Textvorlage verstanden hat, sondern auch, wie gut er über das Verstandene Auskunft geben kann. Zu diskutieren sind auch Fragen der **Reliabilität**: Wie zuverlässig ist z.B. eine Fähigkeit erfasst, wenn zur Überprüfung des globalen Textverständnisses lediglich *ein* „Multiple Choice-Item“ eingesetzt wurde?

Allerdings ist die Einschätzung der Güte von Testaufgaben nicht nur eine psychometrische Angelegenheit, sondern verlangt v.a. hinsichtlich der Inhaltsvalidität fachdidaktische und fachwissenschaftliche Expertise. Man muss kennen und ver-

stehen, 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.

Halten wir also fest: Eine deutschdidaktische Perspektive auf Testaufgaben ist unverzichtbar, wenn es festzulegen gilt, *was* anlässlich einer Aufgabe überprüft – und zuvor natürlich erworben – werden kann.

Aus der Perspektive ihrer spezifischen Domänen (Lesen und Schreiben, Sprechen und Zuhören) gilt es, Wissensvoraussetzungen zu dimensionieren, Aufgabenformate und -arten zu diskutieren und schwierigkeitsgenerierende Potenziale der Gegenstände zu diagnostizieren. Ohne domänenspezifische Expertise wird sich der – künftig zu gewinnende – Zusammenhang von Aufgaben und Kompetenzmodellen kaum herstellen lassen. So haben Behrens/Weirich (2016) bezogen auf den Bereich „Sprechen und Zuhören“ exemplarisch vorgeführt, wie sich die beiden Perspektiven produktiv verbinden lassen

4.3.3 Diskussion

Aufgabenkritik ist immer leichter als die Aufgabengenerierung. Gleichwohl ist Kämper-van den Boogaart zuzustimmen, dass literarische Kompetenzen mit Testfragen nicht hinreichend erfasst werden. Nach Kämper-van den Boogaart (2010) dominieren eindeutig Leseaufgaben in der Dimension „Informationen ermitteln“. „Wenn es überhaupt zu textbezogenem Interpretieren kommt, dann sind z. B. Aufgaben kaum vertreten, bei denen es um das text- und vor allem vorwissensgesteuerte Verknüpfen von über den Text verstreuten Informationen geht“ (Kämper-van den Boogaart 2010, 8). Tatsächlich finden sich in den Tests von VERA 6 relativ wenige Items, die die Etablierung globaler Kohärenz und Sinnzuweisung überprüfen sollen. Dieser Mangel an entsprechenden Aufgaben ist sowohl auf Probleme der Aufgabenkonstruktion zurück zu führen als auch auf das Scheitern in der Erprobung (Tauglichkeitsprüfung).

(1) Welche Probleme zeigen sich im Bereich der Aufgabenkonstruktion?

Grundsätzlich erscheint es schwierig, gute Items für die zuletzt genannten Verstehensebenen zu generieren. Es gibt Items, die zwar auf die Etablierung globaler Kohärenz und auf Sinnzuweisung zielen, aber eben auch ohne Erzeugung eines Textweltmodells richtig gelöst werden können. Wenn das der Fall ist, dann fehlt dem Item die erforderliche Konstrukt-Validität.

Bei Multiple-Choice-Items besteht die Schwierigkeit, plausible und zugleich unzutreffende bzw. nicht hinreichende Distraktoren auf diesen Verstehensebenen zu finden.

Je nach Textbeschaffenheit lassen sich durchaus Kurzantwort-Items generieren, die auf die Pointe eines Textes gerichtet sind und folglich ein Textweltmodell voraussetzen. Die Schwierigkeit ist hier eine doppelte: Der Aufgabenentwickler muss den Text durchschauen und zugleich eine treffende Frage formulieren, die im Rahmen einer Kurzantwort zu bewältigen ist.

Offene Items sind zwar einfach zu formulieren, aber die Erstellung adäquater Kodieranweisungen erweist sich häufig als problematisch. Das gilt vor allem dann, wenn die Erschließung von Textsinn oder Textinterpretation getestet wird. Das hat in erster Linie damit zu tun, dass das Format eine große Lösungsvielfalt begünstigt. Folglich ist es schwierig, zutreffende Lösungen nicht nur auf einem hinreichenden Abstraktionsniveau zu formulieren, sondern auch konkrete Beispiele für richtige und falsche Antworten zu geben. Zu diesem Zweck müssen die Items vorab erprobt werden, um Lösungen aus Lernerperspektive in die Korrekturanleitung aufzunehmen. Sonst besteht die Gefahr, dass Lösungen der Lehrpersonen als Meisterlösung absolut gesetzt werden und andere ebenfalls zutreffende Lösungen gar nicht in den Horizont der Entwickler geraten.

(2) Warum bestehen viele Items die Pilotierung nicht?

Ein Multiple-Choice-Item ist nicht funktionsfähig, wenn z.B. eine oder mehrere Antwortmöglichkeiten ungenutzt bleiben.

Zahlreiche offene Items werden nicht für den Haupttest zugelassen, weil die Korrektoren zu divergierenden Urteilen kommen (fehlende Kodiererübereinstimmung).

Ein dritter Grund für das Ausscheiden eines Items liegt vor, wenn das Item nicht hinreichend differenziert. Hinter dem Differenzierungsanspruch steht die zentrale Annahme: Je höher die Kompetenz, desto höher die Lösungswahrscheinlichkeit. Entspricht ein Item bei der Pilotierung dieser Annahme nicht, dann spricht man von fehlender Differenzierungsfähigkeit.

Die kritische Diskussion der normierten Testaufgaben ist darüber hinaus auch auf eine Reihe von Folgen gerichtet. Gestützt auf angelsächsische Studien resümiert Kämper-van den Boogaart (2010) eine Reihe gravierender Folgen vor allem der Tests, die mit starken Konsequenzen für die Beteiligten verbunden sind. Die Folgen werden auf drei Ebenen wirksam.

Auf der Ebene des Unterrichts führten die Tests und die darauf gerichteten Vorbereitungen zu einer testbezogenen **Curriculumsverengung**. Was nicht getestet wird, werde im Unterricht marginalisiert oder völlig ausgeklammert. Ein Spezialfall solcher Verengung sind **normierte Lektüren** (8), die festlegen, im Hinblick auf welche Zielsetzungen ein literarischer Text zu lesen sei. Dass dabei die **Polyvalenz-Konvention** (8) ausgeschaltet wird, erklärt sich von selbst.