

Inger Petersen, Raja Reble, Jörg Kilian (Hrsg.)

# Texte schreiben in allen Unterrichtsfächern

Textbeurteilung als Grundlage für  
Schreibförderung und Leistungsbewertung

# Sprachliche Bildung

herausgegeben vom  
Mercator-Institut für Sprachförderung  
und Deutsch als Zweitsprache

Band 11

Inger Petersen, Raja Reble, Jörg Kilian (Hrsg.)

# Texte schreiben in allen Unterrichtsfächern

Textbeurteilung als Grundlage  
für Schreibförderung und  
Leistungsbewertung



Waxmann 2024

Münster • New York

### **Bibliografische Informationen 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.

### **Sprachliche Bildung, Band 11**

ISSN 2510-4519

Print-ISBN 978-3-8309-4903-9

E-Book-ISBN 978-3-8309-9903-4

© Waxmann Verlag GmbH, Münster 2024

Steinfurter Straße 555, 48159 Münster

[www.waxmann.com](http://www.waxmann.com)

[info@waxmann.com](mailto:info@waxmann.com)

Umschlaggestaltung: Anne Breitenbach, Münster

Umschlagfoto: © FangXiaNuo | [istockphoto.com](https://www.istockphoto.com)

Satz: Roger Stoddart, Münster

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.  
Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

# Inhalt

*Inger Petersen, Raja Reble & Jörg Kilian*  
Einleitung .....7

## **Textbeurteilung in der schreibdidaktischen Forschung**

*Afra Sturm*  
Lernförderliches Feedback im Bereich Schreiben .....19

*Corinna Peschel*  
Kann man das Beurteilen von (Schüler-)Texten lernen?  
Einblicke in diagnostisches Vorgehen von Lehrkräften. ....39

*Inga Christiana Eckardt & Inger Petersen*  
Diagnostische Prozesse bei der kriteriengeleiteten Analyse von  
lersersprachlichen Texten aus dem Fachunterricht .....61

*Nadja Lindauer*  
Effektivität und Effizienz von holistischen Benchmarkratings .....91

*Jörg Kilian, Lea Vanselow, Frederike Strahl, Thorben Jansen & Jens Möller*  
Zur Bewertung von ideologischem Sprachgebrauch und  
Werteverletzungen in Schülerinnen- und Schülertexten  
Grundlagen und Ergebnisse einer ASSET-G-Sonderstudie ..... 113

## **Textbeurteilung in der fachdidaktischen Forschung**

*Claudia Forkarth & Sabine Manzel*  
Diagnoseinstrumente zur Beurteilung und Förderung von  
Schreibkompetenzen im sozialwissenschaftlichen Fachunterricht ..... 131

*Charlotte Husemann*  
Das Historische Sachurteil auf dem Prüfstand  
Vorstellung eines Stufenmodells als Grundlage für die Beurteilung von  
schriftlichen Sachurteilstexten von Schüler\*innen im Geschichtsunterricht. .... 149

*Christine Enzenbach & Heiko Krabbe*

Welche Rolle spielt die Sprache im Physikunterricht bei der  
Darstellung und Beurteilung von fachlichen Leistungen?  
Didaktische Handlungsfelder beim textsortenbasierten Schreiben ..... 169

*Ulf Abraham & Daniela Matz*

Kreatives und literarisches Schreiben bewerten ..... 199

## **Textbeurteilung in der pädagogisch-psychologischen Forschung**

*Frederike Strahl*

Die diagnostische Kompetenz von Lehrkräften bei der Beurteilung  
von Texten  
Theoretische Grundlagen, Einflussfaktoren, aktuelle Forschungsergebnisse  
und Implikationen für die Praxis ..... 221

*Thorben Jansen, Raja Reble & Jens Möller*

Der Big-Fish-Little-Pond-Effekt bei der Beurteilung  
von Schülertexten ..... 241

*Moti Mathiebe, Joachim Grabowski & Michael Becker-Mrotzek*

Reliabilität bei der Beurteilung von Textqualität ..... 253

Stichwortverzeichnis ..... 277

Autorinnen und Autoren ..... 279

## Einleitung

Das Schreiben von Texten spielt in fast allen Unterrichtsfächern eine große Rolle. In jedem dieser Fächer gibt es zentrale Textgenres, die bestimmten Zielen des Fachunterrichts dienen. Sie werden oft durch bestimmte Operatoren gefordert, die fachspezifisch interpretiert werden und zu den entsprechenden Texten führen sollen (vgl. etwa den Operator *beschreiben* im Fach Biologie und den Operator *beschreiben* im Fach Deutsch). Eine wichtige Bedeutung kommt in den meisten Unterrichtsfächern der kognitiven Funktion des Schreibens zu: Die Schüler\*innen schreiben, um ihr (Arbeits-)Gedächtnis zu entlasten, Informationen untereinander zu vernetzen und neue Erkenntnisse zu gewinnen (Ossner, 1995). Schreiben hilft dabei, komplexe Sachverhalte besser zu erschließen, vorhandenes Wissen zu vertiefen und neues Wissen zu entwickeln. Dieses epistemische Potenzial wird im Unterricht allerdings noch zu wenig genutzt (Thürmann, Pertzel & Schütte, 2015). Die Forschung legt nahe, dass Schüler\*innen im Unterricht zwar wichtige Informationen festhalten, indem sie Notizen machen oder Inhalte kurz zusammenfassen; zusammenhängende Texte werden aber nur selten geschrieben (Decker & Hensel, 2019). Das ist auch insofern problematisch, als es eine „scheinbar triviale, aber dennoch häufig unterschätzte Tatsache ist, dass gerade leistungsschwache Schülerinnen und Schüler ihre Schreibkompetenz nur dann erfolgreich ausbauen können, wenn sie häufig schreiben“ (Marx & Steinhoff, 2017, S. 182). Schüler\*innen benötigen zudem ein regelmäßiges und qualitativ hochwertiges Feedback zu ihren Texten, um ihre Schreibkompetenzen erfolgreich ausbauen zu können (Graham, Hebert & Harris, 2015). Ein weiterer, potenziell problematischer Aspekt kommt hinzu: Schulisch-institutionell überwiegt beim Schreiben eine andere Funktion als die epistemische, nämlich die Funktion des Schreibens zur Leistungsbeurteilung. Dieser Umstand hat in der Forschung bisher jedoch nur wenig Beachtung gefunden. Wird dabei die Produktion zusammenhängender Texte erwartet, dann wird zusammen mit dem Nachweis der jeweils fachspezifischen Kompetenz immer auch die Fähigkeit bewertet, einen funktional angemessenen, adressatenorientierten und verständlichen Text im jeweiligen Fach zu schreiben. Dementsprechend müsste die Entwicklung dieser Fähigkeit aber auch eine stärkere Bedeutung im Fachunterricht bekommen.

Erfreulicherweise hat das schulische Schreiben außerhalb des Deutschunterrichts in den letzten Jahren im wissenschaftlichen Diskurs vermehrte Aufmerksamkeit erfahren. Man könnte sagen, dass der „Weckruf zum Schreiben im Fachunterricht“ (Thürmann et al., 2015) eine gewisse Wirkung entfaltet hat. Es sind im deutschsprachigen Raum verschiedene Publikationen entstanden, die sich u. a. mit folgender Frage beschäftigen: „Welchen Beitrag kann das Schreiben zum erfolgreichen Lernen in den Fächern und verschiedenen Lernbereichen leisten, und wie sind Schreibaktivitäten im Fach zu gestalten, damit sie Lernprozesse wirksam

unterstützen?“ (Schmölzer-Eibinger & Thürmann, 2015, S. 10). In diesem Zusammenhang sind beispielsweise verschiedene fächerspezifische schreibdidaktische Konzepte erarbeitet und Untersuchungen durchgeführt worden (vgl. z. B. Busse, Müller & Siekmann, 2022), die an der Schnittstelle zwischen Sprach- und Fachdidaktik angesiedelt sind. Zudem liegen Studien zur fachsprachlichen Schreibkompetenz in unterschiedlichen Fächern vor. Die Untersuchungen von Roll et al. (2019) im Rahmen des SchriFT-Projektes<sup>1</sup> weisen beispielsweise auf enge Zusammenhänge zwischen fachlicher und fachsprachlicher Textsortenfähigkeit und Fachwissen hin. Zwischen fachspezifischer und fächerübergreifender Textsortenkompetenz, wie sie im Deutschunterricht vermittelt werden sollte, zeigen sich hingegen nur schwache Zusammenhänge (Roll et al., 2019, S. 36), was den Schluss zulässt, dass die fachspezifische Schreibkompetenz nur im Fach selbst ausgebildet werden kann. Diese Befunde ließen sich allerdings in einer weiteren Studie nur teilweise replizieren (Roll et al., 2022).

Studien zur Wirksamkeit von schreibdidaktischen Interventionen führen zu gemischten Ergebnissen: Eine Studie bei Deutschschweizer Gymnasialschüler\*innen (10. und 11. Schuljahr) im Fach Geschichte zeigt, dass eine acht Doppellektionen umfassende Schreibintervention, die sowohl prozess- als auch produktorientiert ausgelegt ist, einen positiven Einfluss auf die Entwicklung der narrativen Kompetenz im Geschichtsunterricht hat (Waldis, Nitsche & Gollin, 2020). Die Interventionsstudien zu den Effekten einer textsortenbasierten Schreibförderung in den Fächern Geschichte, Physik, Politik und Technik (Jahrgangsstufen 7 und 8), die im Rahmen des Projektes SchriFT II (Roll et al., 2022) durchgeführt wurden, ergeben für die Fächer Geschichte, Physik und Politik teilweise geringe Lernzuwächse, jedoch keine signifikanten Interventionseffekte. Lediglich im Fach Technik konnte ein Interventionseffekt für die sprachliche Handlung des Beschreibens festgestellt werden (Roll et al., 2022, S. 18f.). Die Autor\*innen führen dies auf die besonderen Bedingungen einer Feldstudie zurück (Roll et al., 2022, S. 19).

Peschel und Mainzer-Murrenhoff (2019) nehmen in ihrer Untersuchung schreibförderdiagnostische Kompetenzen von Lehramtsstudierenden in den Blick. Sie konstatieren, dass die Studierenden bei der Diagnose und Förderung domänenspezifischen Aspekten zu wenig Beachtung schenken und führen dies einerseits auf mangelndes Fachwissen zurück, andererseits sehen sie eine Verknüpfung mit „der aktuellen Diagnose- und Materiallandschaft, d. h. dem, was den Studierenden an ‚Werkzeug‘ für das förderdiagnostische Handeln zur Verfügung steht“ (Peschel & Mainzer-Murrenhoff, 2019, S. 24). Andere aktuelle Arbeiten nehmen sich der Frage an, wie der Umgang mit authentischen Schüler\*innentexten bereits im Studium die professionelle Kompetenz von Lehramtsstudierenden im Bereich Sprachbildung fördern kann (vgl. Enzenbach, 2022; Petersen & Eckardt, 2023; Schroedler, Enzenbach, Kania & Roll, i. E.).

---

1 Nähere Hinweise zum SchriFT-Projekt (Schreiben im Fachunterricht der Sekundarstufe I unter Einbeziehung des Türkischen) unter <https://www.uni-due.de/schrift/> (08.03.2024).



Untersuchungen zu Fragen der Beurteilung von Texten als Ausgangspunkt für Leistungsbewertung und Förderung in allen Unterrichtsfächern sind allerdings eher rar. Sieht man einmal von den (ebenfalls eher wenigen) Untersuchungen zur Textbewertung im Fach Deutsch ab (vgl. z.B. Birkel, 2003; Abraham, 2010; Steinhoff, 2010; Grausam, 2018), dann sind hier namentlich Untersuchungen aus der Pädagogik und der Pädagogischen Psychologie zu nennen (vgl. z.B. Südkamp, Kaiser & Möller, 2012; Hennes, 2020). Insbesondere psychologisch ausgerichtete Studien machen auf mögliche Urteilsfehler aufmerksam und zeigen beispielsweise den Einfluss von einzelnen sprachlichen Oberflächenmerkmalen auf die Beurteilung von Textqualität (z.B. orthografische Fehler; Jansen, Vögelin, Machts, Keller & Möller, 2021). Im Projekt ASSET-G<sup>2</sup> werden in Kooperation zwischen Deutschdidaktik und Pädagogischer Psychologie insgesamt acht Teilstudien zu unterschiedlichen Einflussfaktoren bei der Beurteilung von deutschen Schüler\*in-entexten durchgeführt, um zu ermitteln, in welchem Ausmaß leistungsrelevante und leistungsirrelevante Merkmale die Textbeurteilung durch angehende und erfahrene Deutschlehrkräfte beeinflussen (vgl. Vanselow, Jansen, Kilian, Strahl & Möller, 2022). Insgesamt fällt auf, dass Fragen der Textbeurteilung und zur Funktion des Schreibens zur Leistungsbeurteilung zwar in der deutschdidaktischen und psychologischen Forschung bearbeitet werden, weniger aber in den Fächern selbst. Eine Ausnahme stellt die Studie von Feser (2019) zur Leistungsbeurteilung im Physikunterricht dar. Der Autor kommt zu dem Ergebnis, dass Lehrkräfte Urteile zu sprachlichen und fachlichen Leistungen miteinander konfundieren. Dieses Ergebnis spricht noch einmal dafür, dass hier ein erheblicher Forschungsbedarf besteht.

Es kann festgehalten werden, dass die Beurteilung von Texten eine zentrale Aufgabe aller Lehrkräfte ist und dass gleichzeitig aus der Perspektive der durchgängigen Sprachbildung noch viele Forschungsdesiderata im Bereich der Textbeurteilung und Schreibförderung in den Unterrichtsfächern bestehen. Dabei erscheint es uns wünschenswert, dass die unterschiedlichen Disziplinen, die im Bereich der Schreibforschung sowie der schreibbezogenen Unterrichts- und Professionalisierungsforschung arbeiten, z.B. Deutschdidaktik und weitere Fachdidaktiken, DaF/DaZ sowie Pädagogische Psychologie, noch stärker kooperieren. Dafür ist es – wie immer in der interdisziplinären Zusammenarbeit – notwendig, sich über die ggf. unterschiedliche Bedeutung der zentralen verwendeten Fachtermini bewusst zu sein. Im Fall des Themas dieses Sammelbandes betrifft diese Mehrdeutigkeit beispielsweise das Lexem der „(Text-)Beurteilung“. Während in der Deutschdidaktik die *Textbeurteilung* in der Regel als „eine verbal geäußerte Bewertung, die hier gegenüber dem Schüler/Schreiber geäußert wird“ (Becker-Mrotzek, 2014, S. 502), gilt, wird der Terminus in bildungs- und erziehungswissenschaftlicher Literatur anders genutzt: Hier zielt er auf „die adäquate Beschreibung und Bewertung einer schulischen Leistung“ (Lissmann, 2010, S. 6), in unserem Fall einer schriftlichen und damit sprachlichen Leistung, ab, die sich

---

2 Nähere Informationen zum ASSET-G-Projekt unter <https://www.asset-g.uni-kiel.de/de/herzlich-willkommen-auf-der-seite-des-asset-g-projektes> (08.03.2024).

zunächst nicht explizit an den\*die Autor\*in des Textes richtet. Beurteilung in diesem Sinne wird in der deutschdidaktischen Literatur eher als (Text-)Bewertung bezeichnet: Damit ist nach Becker-Mrotzek (2014) der „kognitive Akt des Einschätzens“ (S. 502) gemeint. Wir finden es zum einen wichtig, sich über diese Unterschiede im Klaren zu sein, zum anderen erscheint uns der Beurteilungs- und Bewertungsbegriff der Deutschdidaktik noch theoretisch unterspezifiziert (vgl. auch Abraham & Matz, in diesem Band). Am Ende dieses Sammelbandes befindet sich ein Register zu thematisch einschlägigen Termini. Wir sind uns darüber bewusst, dass es sich dabei nur um eine kleine Auswahl an einschlägigen Lexemen handelt. Gleichzeitig erhoffen wir uns aber, dass wir mit dem Register eine Lektüre über die fachdisziplinären Grenzen hinweg und eine Reflexion über die (unterschiedliche) Verwendung der Termini anregen können.

In dieser Publikation sind Beiträge von Vertreter\*innen der Schreibforschung, der Sprach- und weiteren Fachdidaktiken sowie der Pädagogischen Psychologie versammelt, die auf dem Feld der Schreibförderung und der Diagnostik der Schreibkompetenz auf der Basis von Texten in den jeweiligen fachlichen Kontexten forschen. Das Buch führt die verschiedenen Perspektiven auf die Textbeurteilung und Schreibförderung im Fachunterricht – die sprachdidaktische, die fachdidaktische und die psychologische – zusammen, damit sie sich gegenseitig befruchten können. Wir hoffen, dass der Sammelband auf diese Weise Impulse für verstärkte interdisziplinäre Forschungsbemühungen zu Fragen der Textbeurteilung als Grundlage für Schreibförderung und Leistungsbewertung in allen Unterrichtsfächern geben wird. Die Publikation richtet sich somit an alle, die die Schreibkompetenz von Schüler\*innen in allen Unterrichtsfächern erforschen und fördern wollen und Schüler\*innentexte auch als Indikatoren der jeweils erreichten Schreibkompetenz lesen.

Der Sammelband besteht aus drei großen Teilen, die sich jeweils auf das Schreiben und Beurteilen von Texten in fachlichen Kontexten beziehen, allerdings aus unterschiedlichen Perspektiven:

1. Textbeurteilung in der schreibdidaktischen Forschung
2. Textbeurteilung in der fachdidaktischen Forschung
3. Textbeurteilung in der pädagogisch-psychologischen Forschung.

Das erste Kapitel „Textbeurteilung in der schreibdidaktischen Forschung“ enthält verschiedene Beiträge aus der Schreibforschung, deren Erkenntnisse auch Impulse für die fächerübergreifende Schreibdidaktik und deren Erforschung geben können.

*Afra Sturm* eröffnet dieses Kapitel mit einem Beitrag, in dem sie Formen und Funktionen des formativen Feedbacks im Zusammenhang mit dem Überarbeiten von Texten fokussiert. Die Textüberarbeitung hat sich in der schreibdidaktischen Forschung spätestens seit der sog. kognitiven Wende als eine bedeutsame Komponente des Schreibprozesses erwiesen; die Fähigkeit, den eigenen Text kritisch lesen und revidieren zu können, ist eine bedeutsame Komponente der Schreibkompetenz. In aktuellen Bildungsstandards und Lehrplänen zählt die Förderung die-

ser Fähigkeit zum festen Bestand des sprachlichen Lernens und der sprachlichen Bildung, doch spielt die Textüberarbeitung in der schulischen Praxis noch immer eher eine Nebenrolle. Vor diesem Hintergrund werden im Beitrag verschiedene Feedback-Verfahren beschrieben und auf der Grundlage von Befunden der empirischen schreibdidaktischen Forschung in Bezug auf ihre Effektivität bewertet.

*Corinna Peschel* geht in ihrem Beitrag der Frage nach, ob und inwiefern das Beurteilen schriftlicher Texte für (angehende) Lehrpersonen lernbar ist und somit auch verbindlicher Teil der Lehrer\*innenbildung nicht nur im Fach Deutsch sein sollte. Zwei Schwerpunkte fokussieren die Erkundungen zur Beantwortung dieser Frage: die theoretische Fundierung einer prozessorientierten Beurteilung und die empirische Ermittlung der Effektivität der Bewusstmachung förderorientierter formativer Ansätze zur Textbeurteilung bei angehenden Lehrkräften. Die Untersuchungen leisten auch einen Beitrag zur Überwindung einer überwiegend an Oberflächenmerkmalen und mehr oder weniger normativen Textsortenkriterien orientierten Beurteilung schriftlicher Leistungen in allen Fächern. Insgesamt zeigt sich die Notwendigkeit, Studierenden ein präzises Beschreibungsinstrumentarium für Textqualität zu vermitteln.

Das Analyseraster, mit dem Studierende in der explorativen Studie von *Inga Eckardt* und *Inger Petersen* arbeiten, setzt genau hier an. Grundlegend ist die Idee, dass Lehrkräfte nicht nur über Wissen zu fachlichen, sondern auch sprachlichen Kriterien zur Textbewertung verfügen sollten. In der präsentierten explorativen Studie werden sprachdiagnostische Prozesse untersucht, die bei Lehramtsstudierenden gesellschaftswissenschaftlicher Fächer bei der Arbeit mit einem Analyseraster zur sprachlichen Qualität von Texten ablaufen. Dabei haben drei Studierendenpaare (N=6) authentische Schüler\*innentexte aus dem Fachunterricht Geschichte, Weltkunde und Wirtschaft/Politik analysiert. Die Gespräche der Studierenden beim gemeinsamen Bearbeiten des Analyserasters wurden audiografiert sowie transkribiert. Anschließend wurden mittels einer qualitativen Inhaltsanalyse verschiedene epistemische Aktivitäten innerhalb der kognitiven Prozesse des Wahrnehmens und Interpretierens rekonstruiert. Die Ergebnisse zeigen, dass das sprachdiagnostische Handeln genau wie diagnostisches Handeln allgemein immer auf einem Vergleich aufbaut. Dies unterstreicht nochmals die Bedeutung von sprachlichen Erwartungshorizonten bei der Textbeurteilung und wirft gleichzeitig die Frage auf, wie (angehende) Lehrkräfte nichtphilologischer Unterrichtsfächer dazu befähigt werden können, sich ihrer sprachlichen Erwartungen an die Schüler\*innen bewusst(er) zu werden.

Das holistische Benchmarkrating zählt zu den Verfahren der Textbeurteilung, die die komplexe Aufgabe der Bewertung von Texten erleichtern und zu einheitlichen Beurteilungen im Sinne der Gütekriterien führen sollen. Im deutschsprachigen Raum ist das holistische Benchmarkrating bislang wenig verbreitet. *Nadja Lindauer* setzt in ihrem Beitrag daher dazu an, das holistische Benchmarkrating aus der Perspektive der Forschung zu beschreiben und in Bezug auf seine Effektivität und Effizienz zu bewerten. Dazu werden Studien referiert, in denen die Validität und Reliabilität – und somit die Effektivität – des holistischen Bench-

markratings belegt werden, sofern die Textqualität insgesamt oder auf den Gesamttext bezogene Textmerkmale im Fokus der Bewertung stehen. In Bezug auf die Effizienz wird zwar ein relativ höherer Aufwand im Zuge der Konstruktion von Ratingskalen festgestellt, der indes dadurch wieder aufgehoben wird, dass die für die jeweiligen Textgenres entwickelten Skalen wiederholt einsetzbar sind.

Jörg Kilian, Lea Vanselow, Frederike Strahl, Thorben Jansen und Jens Möller präsentieren Grundlagen, Methoden und Befunde einer kleinen Sonderstudie zum ASSET-G-Projekt zur Bewertung von ideologischem Sprachgebrauch und Werteverletzungen in Schülerinnen\*texten. Im deutschdidaktischen Diskurs ist diese Fragestellung bislang kaum in den Blick genommen worden. Der Beitrag referiert erste Befunde aus einer pilotierenden Beurteilung (inkl. Feedback) von Modelltexten durch angehende Deutschlehrer\*innen (N=23) sowie deren anschließende Befragung in Form eines Leitfadeninterviews. Aus den Befunden können Bewertungskriterien abgeleitet werden, die auch für die fachintegrierte Sprachbildung in anderen Fächern (z. B. Geschichte, Wirtschaft/Politik) dazu beitragen können, die Spannung zwischen dem Neutralitätsgebot und dem Gebot der Werteerziehung zumindest zu mildern.

Auch die Beiträge in dem Kapitel „Textbeurteilung in der fachdidaktischen Forschung“ beschäftigen sich mit der Frage, mit welchen Kriterien der fachspezifische sprachliche Erwartungshorizont an Texte operationalisiert werden kann. Vorhandene Kriterienkataloge stammen i. d. R. aus der Deutschdidaktik und bilden nicht die besonderen bildungs- und fachsprachlichen sowie fachlichen Anforderungen ab, durch die sich die verschiedenen Textgenres in den Unterrichtsfächern auszeichnen.

Claudia Forkath und Sabine Manzel setzen sich mit Schreibkompetenzen im Kontext von politischen Urteilen im sozialwissenschaftlichen Fachunterricht auseinander und gehen davon aus, dass Lehrkräfte für die Diagnose von schriftlichen Textprodukten im Fachunterricht Politik sowohl Textsortenkompetenz zu politischen Urteilen als auch transparente Beurteilungsmaßstäbe benötigen. Diese müssen sowohl sprachliche als auch fachliche Dimensionen umfassen. Die Autorinnen präsentieren ein Diagnoseinstrument für die schulische Praxis, das auf einem empirischen Messinstrument zur Schreibkompetenz beruht und aus den Dimensionen *Textstrukturierung*, *Fachinhalt*, *fachliche Begründung*, *Kohäsion* sowie *Fachsprache* besteht. Damit bietet es konkrete Anknüpfungspunkte für eine individuelle Textbeurteilung und -überarbeitung.

Auch Charlotte Husemann setzt sich zum Ziel, ein Instrument zur fach(sprach)-lichen Bewertung von schriftlichen Urteilstexten zu entwickeln, allerdings für den Geschichtsunterricht. Sie unterbreitet anhand empirischer Befunde einen Vorschlag für ein Stufenmodell zur Beurteilung von schriftlichen Sachurteilstexten. Zunächst wird auf der Grundlage einer Stichprobe von N=152 Gesamtschüler\*innen der 7. und 8. Jahrgangsstufe die theoretische Konzeptionierung der Textsorte überprüft. Anschließend präsentiert sie ein fünfstufiges Modell, das die Operatoren *Beschreiben*, *Erklären*, *Begründen* sowie *übergreifende sprachliche Mittel* und *Textstruktur & Kohärenz* als Kategorien ansetzt und die Beurteilung Historischer

Sachurteile erlaubt. Dabei kann die Beschreibung des historischen Sachverhalts, dessen kausale Einbettung (Erklärung) und die abschließende Begründung der Einschätzung des Sachverhalts beurteilt werden.

*Christine Enzenbach* und *Heiko Krabbe* stellen sich in ihrem Beitrag die ganz grundsätzliche Frage, welche Rolle die Sprache beim Lernen und Darstellen von Physik hat. Sie arbeiten heraus, wie bedeutsam die Beherrschung eines bildungs- und fachsprachlichen und damit konzeptionell schriftlichen Registers ist, und dass regelmäßige Schreibenanlässe deshalb auch in den Physikunterricht Eingang finden sollten. Als ein fruchtbarer Ansatz erscheint hier der textsortenbasierte Lehr-Lernzyklus. Da in den Textsorten fachliche Strukturen und sprachliche Mittel funktional aufeinander bezogen seien, eigneten sie sich gut für die Integration der Sprachbildung in den Physikunterricht. Die im Lehr-Lernzyklus erarbeiteten und eingeübten Textsortenmerkmale können schließlich auch für die Textbeurteilung genutzt werden, wie die Autor\*innen am Beispiel des Versuchsprotokolls aufzeigen.

Das Teilkapitel schließt mit einem Beitrag aus der Deutschdidaktik. *Ulf Abraham* und *Daniela Matz* untersuchen die Bewertung von Texten, die auf der Basis von Aufgaben des kreativen und literarischen Schreibens verfasst wurden. In diesem Bereich schulischer Textproduktion sind besondere Kompetenzen gefordert und sollen durch die Bewertung festgestellt werden. Da die entstehenden Texte sehr weitgehend in die Arbeit eines anderen Lernbereichs (mit Texten und Medien umgehen) eingebunden sowie zudem oft besonders subjektorientiert sind, bedarf die Bewertung besonderer Verfahren und Kriterien – ähnlich wie die Bewertung von Leistungen im Kunstunterricht. Die deutschdidaktische Forschung hat sich Fragen der Bewertung von Texten, die als Lösung von Aufgaben des kreativen und literarischen Schreibens entstehen, bislang kaum gewidmet. Abraham und Matz setzen dazu an, die Ansicht zu revidieren, solche Texte seien mangels objektiver Maßstäbe nicht zu bewerten oder zu benoten. Sie plädieren für eine implizite Bewertung, stellen deren Prinzipien dar und entfalten Möglichkeiten für deren Umsetzung. Ergänzend dazu erörtern sie Bedingungen und Kriterien einer ggf. notwendigen expliziten Bewertung.

Das letzte Kapitel „Textbeurteilung in der pädagogisch-psychologischen Forschung“ wird von *Frederike Strahl* mit einem Beitrag zur diagnostischen Kompetenz von Lehrkräften bei der Beurteilung von Texten eröffnet. Sie bietet in ihrem Beitrag aus pädagogisch-psychologischer Perspektive einen Überblick über die theoretischen Grundlagen, Einflussfaktoren und aktuelle Forschungsergebnisse zur diagnostischen Kompetenz von Lehrkräften bei der Beurteilung von Schüler\*innentexten. Dafür stellt sie zunächst zentrale Modelle vor, die die Determinanten und Vorgänge der Urteilsgenauigkeit beschreiben und erörtert deren Anpassung an die Lehrer\*innenbildung und den schulischen Kontext. Damit veranschaulicht sie die vielfältigen Faktoren und Prozesse, die die Urteilsgenauigkeit von Lehrkräften beeinflussen können. Weiterhin referiert sie den aktuellen Forschungsstand zur diagnostischen Kompetenz von Lehrkräften bei der Beurteilung von Schüler\*innentexten. Sie hebt bedeutende Modelle und Studien hervor, die

insbesondere in den Fachdidaktiken mehr Beachtung verdienen. Ihr Beitrag verdeutlicht die Fehleranfälligkeit bei der Textbeurteilung und die Vielfalt der Faktoren, die Lehrer\*innenurteile beeinflussen können. Sie plädiert daher dafür, die diagnostische Kompetenz von Lehrkräften in der Lehrer\*innenbildung zu stärken, um eine angemessene Beurteilung von Schüler\*innentexten in allen Unterrichtsfächern zu gewährleisten.

*Thorben Jansen, Raja Reble und Jens Möller* untersuchen den Big-Fish-Little-Pond-Effekt bei der Beurteilung von Schüler\*innentexten. Lehrkräfte sollen Leistungen von Schüler\*innen unabhängig von den Leistungen ihrer Mitschüler\*innen beurteilen. Empirisch zeigt sich, dass, wenn Lehrkräfte mehrere Leistungen von Schüler\*innen zusammenfassen, wie bei der Vergabe von Schuljahresnoten oder der mündlichen Noten nach einer Schulstunde, sie die Leistungen nicht unabhängig vom Leistungsniveau der Klasse beurteilen. Die Autor\*innen fragen sich daher, ob ein Referenzrahmen-Effekt auch auftritt, wenn ganz konkrete Leistungen wie Schüler\*innentexte beurteilt werden. Ihre experimentelle Studie untersucht, ob ein Schülertext von angehenden Lehrkräften aller Fächer (N=64) unterschiedlich beurteilt wird, je nachdem, ob die angehenden Lehrkräfte zusätzlich zwei Vergleichstexte mit hoher oder niedriger Qualität beurteilen. Jeder Text wurde auf einer holistischen und einer analytischen Skala beurteilt. Die Ergebnisse zeigen, dass Lehrkräfte einen Schülertext auf beiden Skalen positiver (negativer) beurteilen, wenn die Textqualität der Vergleichstexte niedrig (hoch) ist.

*Moti Mathiebe, Joachim Grabowski und Michael Becker-Mrotzek* behandeln in ihrem Beitrag die Herausforderungen der zuverlässigen Beurteilung von Textqualität bei Schüler\*innentexten. Dabei liegt der Fokus auf der Reliabilität als Maß für die Zuverlässigkeit einer Messung. Die Autor\*innen beschreiben typische Faktoren, die die Reliabilität von Textbeurteilungen beeinflussen. Weiterhin nutzen sie die Daten eines mehrphasigen Forschungsprojektes zur Förderung der Schreibkompetenz in der Sekundarstufe I, um die Ergebnisse verschiedener Ratingverfahren von Schüler\*innentexten darzustellen. Die Textqualität wird anhand analytischer und holistischer Verfahren sowie eines praxisnahen Ratings beurteilt. Die Ergebnisse verdeutlichen, dass selbst standardisierte Bewertungskriterien und intensive Schulungen von Lehrkräften keine ausreichend hohe Übereinstimmung bei der Beurteilung von Schüler\*innentexten gewährleisten. Die Autor\*innen diskutieren typische Probleme und Herausforderungen bei der Sicherung der Reliabilität und ermutigen dazu, neue Ansätze zur zuverlässigen Beurteilung der Textqualität zu erforschen.

Insgesamt zeigen alle drei Beiträge dieses Kapitels deutlich, wie fehleranfällig und komplex Textbeurteilungsprozesse sind. Dies unterstreicht noch einmal die Relevanz von interdisziplinären Forschungsbemühungen und einer Professionalisierung der Lehrkräfte im Bereich textbezogener diagnostischer Kompetenz. Wir hoffen, mit unserem Sammelband einen Beitrag zu diesem Forschungsfeld zu leisten.

Wir möchten an dieser Stelle den Gutachtenden aus dem anonymen Peer-Review-Verfahren unseren herzlichsten Dank aussprechen. Sie haben wesentlich

zum Zustandekommen dieses Sammelbandes und zu seiner Qualität beigetragen. Zudem wurden wir bei der Redaktion und Fertigstellung des Sammelbandes tatkräftig von Madita Holland und insbesondere Inga-Sophie Jacobsen unterstützt. Auch ihnen gebührt unser Dank. Nicht zuletzt danken wir dem Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache der Universität zu Köln und den Reihenherausgebern Michael Becker-Mrotzek und Hans-Joachim Roth für die Aufnahme in die Reihe „Sprachliche Bildung“.

## Literatur

- Abraham, Ulf (2010). Bewertung produktiv-kreativer schriftlicher Texte. In Michael Kämper-van den Boogaart, Kaspar H. Spinner (Hrsg.), *Lese- und Literaturunterricht* (Teil 3) (S. 89–115). Baltmannsweiler: Schneider Hohengehren.
- Abraham, Ulf & Matz, Daniela (2024). Kreatives und literarisches Schreiben bewerten. In Inger Petersen, Raja Reble & Jörg Kilian (Hrsg.), *Texte schreiben in allen Unterrichtsfächern. Textbeurteilung als Grundlage für Schreibförderung und Leistungsbewertung* (S. 199–219). Münster/New York: Waxmann.
- Becker-Mrotzek, Michael (2014). Schreibleistungen bewerten und beurteilen. In Helmut Feilke & Thorsten Pohl (Hrsg.), *Schriftlicher Sprachgebrauch und Texte verfassen* (S. 501–513). Baltmannsweiler: Schneider Hohengehren.
- Birkel, Peter (2003). Aufsatzbeurteilung – ein altes Problem neu untersucht. *Didaktik Deutsch*, 15, 46–61.
- Busse, Vera; Müller, Nora & Siekmann, Lea (Hrsg.). (2022). *Schreiben fachübergreifend fördern*. Hannover: Kallmeyer in Verbindung mit Klett.
- Decker, Lena & Hensel, Sonja (2019). Zum Stellenwert des Schreibens im Fachunterricht der gymnasialen Oberstufe – empirische Befunde und schreibdidaktische Konsequenzen. In Lena Decker & Kirsten Schindler (Hrsg.), *Von (Erst- und Zweit-)Spracherwerb bis zu (ein- und mehrsprachigen) Textkompetenzen* (Kölner Beiträge zur Sprachdidaktik, Bd. 13) (S. 49–62). Duisburg: Gilles & Francke.
- Enzenbach, Christine (2022). Authentische Lerner\*innentexte aufgabengeleitet nutzen – zum hochschuldidaktischen Potenzial von Lerner\*innentextkorpora im DaZ-Modul. In Katja F. Cantone, Erkan Gürsoy, Ina Lammers & Heike Roll (Hrsg.), *Fachorientierte Sprachbildung und sprachliche Vielfalt in der Lehrkräftebildung*. Hochschuldidaktische Formate an der Universität Duisburg-Essen (S. 77–90). Münster: Waxmann.
- Feser, Markus S. (2019). *Physiklehrkräfte korrigieren Schülertexte. Eine Explorationsstudie zur fachlich-konzeptuellen und sprachlichen Leistungsfeststellung und -beurteilung im Physikunterricht*. Berlin: Logos.
- Graham, Steve; Hebert, Michael, & Harris, Karen R. (2015). Formative assessment and writing: A metaanalysis. *The Elementary School Journal*, 115(4), 523–547. <https://doi.org/10.1086/681947>
- Grausam, Nina Caroline (2018). *Diagnosekompetenz von Lehrpersonen als Voraussetzung individueller Förderung im Bereich „Texte schreiben“: Eine empirische Studie am Beispiel einer neu eingeführten integrierten Schulform*. Münster/New York: Waxmann.
- Hennes, Ann-Kathrin (2020). *Schreibprodukte bewerten – Die Rolle der Expertise bei der Bewertung der Textproduktionskompetenz*. Dissertation, Universität zu Köln. Verfügbar unter: <https://d-nb.info/1215837100/34>.

- Jansen, Thorben; Vögelin, Christina; Machts, Nils; Keller, Stefan Daniel & Möller, Jens (2021). Don't Just Judge the Spelling! The Influence of Spelling on Assessing Second-Language Student Essays. *Frontline Learning Research*, 9(1), 44–65. <https://doi.org/10.14786/flr.v9i1.541>
- Lissmann, Urban (2010). Leistungsbeurteilung gestern, heute und morgen. In Günther L. Huber (Hrsg.), *Enzyklopädie Erziehungswissenschaft Online*, Fachgebiet Pädagogische Psychologie. Weinheim: Juventa.
- Marx, Nicole & Steinhoff, Torsten (2017). Schreiben von Schülerinnen und Schülern mit nichtdeutscher Familiensprache. In Michael Becker-Mrotzek & Hans-Joachim Roth (Hrsg.), *Sprachliche Bildung – Grundlagen und Handlungsfelder* (S. 175–183). Münster/New York: Waxmann.
- Ossner, Jakob (1995). Prozessorientierte Schreibdidaktik in Lehrplänen. In Jürgen Baurmann & Rüdiger Weingarten (Hrsg.), *Schreiben. Prozesse, Prozeduren und Produkte* (S. 29–50). Opladen: Westdeutscher Verlag. [https://doi.org/10.1007/978-3-322-97050-3\\_2](https://doi.org/10.1007/978-3-322-97050-3_2)
- Peschel, Corinna & Mainzer-Murrenhoff, Mirka (2019). Domänenspezifische Diagnose und Förderung des Schreibens: Was brauchen angehende Lehrkräfte? In Una Dirks (Hrsg.), *Sprachsensible DaF-/DaZ-/DaM-Bildungsräume: Sprech- & Textformen im Fokus* (S. 61–90). Philipps-Universität Marburg. <https://doi.org/10.17192/es2019.0051>
- Petersen, Inger & Eckardt, Inga Christiana (2023). Studierende analysieren Schüler\*innentexte in fächerübergreifenden Lehrveranstaltungen zum Thema Sprachbildung. Ein Analyseraster für lernersprachliche Texte aus dem Fachunterricht. In Melanie Greinert, Thilo Kleickmann & Ilka Parchmann (Hrsg.), *Lehramt mit Perspektive: Ansätze, Methoden und Forschungsergebnisse einer zukunftsorientierten Lehrkräftebildung* (S. 205–234). Münster/New York: Waxmann.
- Roll, Heike; Bernhardt, Markus; Enzenbach, Christine; Fischer, Hans E.; Forkarth, Claudia; Gürsoy, Erkan; Krabbe, Heiko; Lang, Martin; Manzel, Sabine & Uluçam-Wegmann, İşil (Hrsg.) (2022). *Schreibförderung im Fachunterricht der Sekundarstufe I: Interventionsstudien zu Textsorten in den Fächern Geschichte, Physik, Technik, Politik, Deutsch und Türkisch*. Münster/New York: Waxmann. <https://doi.org/10.31244/9783830994893>
- Roll, Heike; Bernhardt, Markus; Enzenbach, Christine; Fischer, Hans E.; Gürsoy, Erkan; Krabbe, Heiko; Lang, Martin; Manzel, Sabine & Uluçam-Wegmann, İşil (2019). *Schreiben im Fachunterricht der Sekundarstufe I unter Einbeziehung des Türkischen. Empirische Befunde aus den Fächern Geschichte, Physik, Technik, Politik, Deutsch und Türkisch*. Münster/New York: Waxmann.
- Schmölzer-Eibinger, Sabine & Thürmann, Eike (2015). Schreiben als Medium des Lernens: Einleitung. In Sabine Schmölzer-Eibinger & Eike Thürmann (Hrsg.), *Schreiben als Medium des Lernens. Kompetenzentwicklung durch Schreiben im Fachunterricht* (S. 9–15). Münster/New York: Waxmann.
- Schroedler, Tobias; Enzenbach, Christine; Kania, Thomas M. & Roll, Heike (i. E.). Authentische Lerner\*innentexte als Gegenstand fallorientierter Hochschullehre zur Vorbereitung angehender Lehrkräfte für den Umgang mit DaZ und Mehrsprachigkeit im Fachunterricht. *HLZ*.
- Steinhoff, Thorsten (2010). Differenzierte Schülertextbeurteilung: Entwicklungs-, Text- und Situierungsdimension. In Thorsten Pohl & Thorsten Steinhoff (Hrsg.), *Textformen als Lernformen* (S. 257–280). Duisburg: Gilles & Francke. Verfügbar unter: [http://koebes.phil-fak.uni-koeln.de/sites/koebes/user\\_upload/koebes\\_07\\_2010.pdf](http://koebes.phil-fak.uni-koeln.de/sites/koebes/user_upload/koebes_07_2010.pdf).



- Südkamp, Anna; Kaiser, Johannes & Möller, Jens (2012). Accuracy of Teachers' Judgements of Students' Academic Achievement: A Meta-Analysis. *Journal of Educational Psychology*, 104(3), 743–762. <https://doi.org/10.1037/a0027627>
- Thürmann, Eike; Pertzel, Eva & Schütte, Anna Ulrike (2015). Der schlafende Riese: Versuch eines Weckrufs zum Schreiben im Fachunterricht. In Sabine Schmölzer-Eibinger & Eike Thürmann (Hrsg.), *Schreiben als Medium des Lernens. Kompetenzentwicklung durch Schreiben im Fachunterricht* (S. 17–45). Münster/New York: Waxmann.
- Vanselow, Lea; Jansen, Thorben; Kilian, Jörg; Strahl, Frederike & Möller, Jens (2022). Gerecht? Zu streng? Zu mild? Das ASSET-G-Projekt zur Erforschung von Einflussfaktoren bei der Bewertung schriftlicher Leistungen im Deutschunterricht. *Der Deutschunterricht*, 74(4), 89–95.
- Waldis, Monika; Nitsche, Martin & Gollin, Kristine (2020). Schülerinnen und Schüler schreiben Geschichte – Eine Interventionsstudie an Deutschschweizer Gymnasien. *Zeitschrift für Geschichtsdidaktik*, 19, 90–108. <https://doi.org/10.13109/zfgd.2020.19.1.90>



## Lernförderliches Feedback im Bereich Schreiben

### 1. Einleitung

Die Fähigkeit, einen eigenen Text zu überarbeiten, differiert zwischen versierten Schreibenden und Novizen bzw. Novizinnen (MacArthur, 2016): So neigen insbesondere Schüler\*innen mit geringen Schreibkompetenzen dazu, hierarchiehohe Prozesse, zu denen das Überarbeiten zählt, wegzulassen;<sup>1</sup> gleichzeitig verfügen sie oftmals über geringes Schreibwissen<sup>2</sup> oder über limitierte Ressourcen zur Steuerung ihres Schreibprozesses, was sich auch in den Schreibleistungen niederschlägt (Graham & Harris, 2000; Lindauer, 2021). Leistungsstudien wie DESI (Neumann & Lehmann, 2008) für die 9. Jahrgangsstufe in Deutschland oder Checks S2 für die 8. Jahrgangsstufe im Bildungsraum Nordwestschweiz (Institut für Bildungsevaluation, 2020) belegen einen recht hohen Anteil an Schüler\*innen, die gegen Ende Sekundarstufe I über geringe Schreibkompetenzen verfügen und damit auch nur in einem geringen Maß die Fähigkeit haben, Texte zu überarbeiten.

Gerade das Vermitteln von Überarbeitungskompetenzen bietet deshalb wichtige Ansatzpunkte in folgender Hinsicht:

- 1) Der Bereich Überarbeiten kann dazu dienen, die Schüler\*innen zu effektivem Schreiben anzuleiten, und zwar nicht nur bezogen auf einen einzelnen konkreten Text, sondern auch mit Blick auf künftige komplexe bzw. herausfordernde Schreibaufgaben (MacArthur, 2016, S. 272).
- 2) Das Überarbeiten umfasst nicht nur die eigentlichen Revisionshandlungen, sondern auch das Evaluieren eines Textes. Damit lässt sich der Umstand nutzen, dass Lesen und Schreiben wesentliche Aspekte teilen (u. a. Philippakos & Graham, 2022). So wirkt sich bspw. die Vermittlung von Genre-Merkmalen und Textstrukturen sowohl positiv auf das Textverstehen wie auch auf das Verfassen eigener Texte aus. Letzteres kann das Evaluieren guter und weniger guter Beispieltexte umfassen (Philippakos, 2022).
- 3) Eine explizite Vermittlung von Strategien (inkl. Überarbeitungsstrategien) beinhaltet immer auch Komponenten, die schreibbezogenes Wissen aufbauen oder erweitern (Sturm, 2022, S. 100f.).

---

1 Hierarchiehohe Prozesse im Schreiben zeichnen sich dadurch aus, dass sie der Reflexion zugänglich sind, während hierarchieniedrige weitgehend automatisiert ablaufen (Olive, 2012).

2 Harris, Graham, Brindle & Sandmel (2009) zählen zum Schreibwissen insbesondere *deklaratives Wissen* über sich als Schreiber\*in, über aufgabenspezifische Anforderungen (inkl. Textsortenwissen) sowie Wissen über Schreibstrategien, *prozedurales Wissen* – wie eine Schreibstrategie auszuführen oder ein bestimmtes Ziel zu erreichen ist – sowie *konditionales Wissen* zur Frage, wann, wie oder warum auf deklaratives oder prozedurales Wissen zurückgegriffen werden sollte.

Studien, die sich dem Ansatz 3) zuordnen lassen, erfassen selten das schreibbezogene Wissen der Schüler\*innen. Zu nennen sind hier Amir, Atkin und Rijlaarsdam (2021), die Schreibwissen in Form von Selbsteinschätzungen erfragten, oder Knott (2023), die Leitfaden-Interviews einsetzte. In der Regel fokussieren Studien zum Schreibwissen jedoch die Schreibentwicklung und die Zusammenhänge mit den Schreibleistungen (vgl. Sommer, 2020).

Damit Schüler\*innen bzw. Schreibnoviz\*innen einen Überarbeitungsbedarf ihrer Texte erkennen können, sind sie auf Feedback angewiesen. Entsprechend kommt der Wahl geeigneter und wirksamer Feedback-Verfahren ein hohes Gewicht zu, zumal die bloße Aufforderung an die Schüler\*innen, einen Text zu überarbeiten, mit  $d = .19$  keine nennenswerte Wirkung zeigt (Rijlaarsdam et al., 2008, S. 55).

In diesem Beitrag steht entsprechend Feedback im Zusammenhang mit dem Überarbeiten von Texten im Vordergrund, und zwar vor allem bezogen auf den Ansatzpunkt 1). Solches Feedback stellt im Kern formatives Feedback dar: In Kapitel 2 wird deshalb kurz auf die Unterscheidung von summativem und formativem Feedback in der Domäne Schreiben eingegangen. Kapitel 3 geht der Frage nach, welchen Stellenwert das Überarbeiten mit Feedback curricular, in Lehrmitteln sowie im Unterricht hat. Darauf aufbauend bietet Kapitel 4 einen Überblick über verschiedene Feedback-Formen: Dabei werden insbesondere die damit verbundenen instruktionalen Prinzipien herausgearbeitet. Die meisten Studien, die hier beigezogen werden, fokussieren das Schreiben im Fach ‚Schulsprache‘: Die Ergebnisse und Prinzipien sind aber auf andere Fächer übertragbar, sofern es um (fachspezifische) Texte geht. Schreibendes Lernen, das zur Unterstützung des fachlichen Lernens eingesetzt wird, wird hier nicht berücksichtigt. Der Beitrag schließt mit einer Diskussion.

## 2. Summatives vs. formatives Feedback und deren Wirkung

Bei *summativem Feedback* geht es weniger um die Frage, welche passende Fördermaßnahme als Nächstes den Lernenden angeboten werden soll, sondern vielmehr soll rückblickend überprüft werden, inwiefern ein bestimmtes Lernziel erreicht worden ist. Entsprechend stellen summative Beurteilungssituationen im Kern Leistungssituationen dar, in denen Schüler\*innen sinnvollerweise Fehler möglichst vermeiden, um das Risiko eines Misserfolgs zu minimieren (Weinert, 1999).

Im Bereich Schreiben kann das kriterienorientierte Beurteilen von Lerner-texten, das in der Regel durch die Lehrperson erfolgt, als prototypisch für summatives Feedback bezeichnet werden. Kriterienraster, die weder genre- oder textsortenspezifisch sind noch Deskriptoren anführen, weisen einen (zu) großen Interpretationsspielraum auf, sodass nicht davon ausgegangen werden kann, dass es sich um zuverlässige Instrumente handelt (vgl. bspw. Graham, McKeown, Kihara & Harris, 2012). Kriterienraster mit geringerem Interpretationsspielraum, die zum einen entwicklungs sensitiv angelegt sind und zum anderen empirisch

validiert wurden, stoßen bei Lehrpersonen auf hohe Akzeptanz, gehen aber mit hohem Aufwand sowie hohen Kosten einher (Skar, Jølle & Aasen, 2020).

*Formatives Feedback* ist anders als summatives Feedback auf die Lernförderung ausgerichtet: „Fehler“ können in solchen Lernsituationen mit einem Erkenntnisgewinn verbunden sein, da sie wertvolle Informationen für den weiteren Lernweg beinhalten. Lernende erachten Feedback als besonders hilfreich, wenn es in den Lernprozess integriert und so informativ ist, dass es zur Verringerung des Unterschieds zwischen aktueller und künftig zu erreichender Lernleistung genutzt werden kann. Ist Feedback vage oder negativ formuliert, zeigt es darüber hinaus keine konkreten Handlungsmöglichkeiten auf oder kommt es zu spät, ist es für Lernende nicht oder zu wenig hilfreich (Brooks, Huang, Hattie, Carroll & Burton, 2019).

Auf die Domäne Schreiben bezogen gilt denn entsprechend auch, dass formatives Feedback in den Textproduktionsprozess integriert und im Text verankert sein sowie dialogisch erfolgen sollte (Cho & MacArthur, 2010). Hinzu kommt, dass wirksames Feedback im Schreiben nicht Oberflächenmerkmale – Handschrift oder Sprachformales – und auch nicht nur lokale Aspekte auf Wort- und Satzebene fokussieren, sondern in erster Linie auf der eigentlichen Textebene ansetzen und damit Aspekte wie Schreibfunktion, Struktur, Adressatenorientierung etc. in den Blick nehmen sollte (vgl. den Überblick in Sturm, Lindauer & Sommer, 2018).

Formatives Feedback beinhaltet des Weiteren Informationen, die auf verschiedenen Ebenen ansetzen können: Feedback wirkt sich allgemein positiv aus, wenn es auf der Ebene der Aufgabe, des Prozesses oder der Selbstregulation ansetzt. Die positive Wirkung lässt sich Hattie und Timperley (2007, S. 90ff.) zufolge erhöhen, wenn diese drei Ebenen kombiniert werden (vgl. dazu auch Kap. 4.4). Feedback auf der persönlichen Ebene (Bsp.: „Du hättest dir mehr Mühe geben können“) zeigt jedoch keine Wirkung auf die Lernleistung.

Meta-Analysen wie jene von Graham, Harris und Hebert (2011, S. 17) oder Koster, Tribushinina, de Jong und van den Bergh (2015, S. 263) berichten für formatives Feedback einen positiven Effekt von  $d = .77$  bzw.  $d = .88$  auf die Schreibleistung.

### 3. Überarbeiten und Feedback im Schreibunterricht

Baurmann und Ludwig (1996) halten fest, dass das Überarbeiten im Aufsatzunterricht bis in die 90er-Jahre keine Rolle gespielt habe, dass aber mit den neuen Lehrplänen eine Umorientierung festzustellen sei, die sich nicht mehr mit dem Einüben bestimmter Aufsatzformen begnügt:

Die Schüler und Schülerinnen sollen schreiben lernen, nicht nach Mustern Texte aufsetzen, und zum Schreiben gehört ganz wesentlich auch die Überarbeitung des Geschriebenen. (Baurmann & Ludwig, 1996, S. 13)

Fast 30 Jahre später ist dennoch davon auszugehen, dass Überarbeiten als Ansatzpunkt für einen effektiven Schreibunterricht und damit auch Feedback-Verfahren insgesamt noch wenig genutzt werden (vgl. Graham, 2019; Rietdijk, van Weijen, Janssen, van Den Bergh & Rijlaarsdam, 2018). Zwei Hauptgründe lassen sich anführen, die im Folgenden genauer ausgeführt werden:

- a) Das Überarbeiten ist nicht immer curricular eingebettet.
- b) Es werden Förderansätze verwendet bzw. in Lehrmitteln angeboten, die wenig wirksam sind bzw. die das Potenzial wirksamer Ansätze noch nicht ausreichend nutzen.

Sowohl die Bildungsstandards in Deutschland (bspw. KMK, 2022a, S. 14ff., für die Primarstufe, sowie KMK, 2022b, S. 23ff., für ESA/MSA) als auch der Schweizer Lehrplan 21 (D-EDK, 2016; für alle Jahrgangsstufen) stärken einen prozessorientierten Zugang, der den Schreibprozess in die Makroprozesse Planen, Formulieren und Überarbeiten gliedert. Curriculare Vorgaben in den deutschen Bundesländern sparen jedoch den Makroprozess Überarbeiten gern aus oder legen den Schwerpunkt auf das Planen (Müller, Lindefeld & Busse, 2021). Entsprechend ist damit zu rechnen, dass im Schreibunterricht das Überarbeiten und damit auch Feedback-Verfahren eher eine geringe Rolle spielen.

Dies stellt sich für die Schweiz etwas anders dar: Zum einen trennt der Schweizer Lehrplan 21 das Überarbeiten weiter in inhaltliches und sprachformales Überarbeiten auf (D-EDK, 2016, D.4E–F), zum anderen präzisiert der Lehrplan 21 das inhaltliche Überarbeiten auch mit Blick auf prozedurale Aspekte, indem er festhält, dass die Schüler\*innen ihren Text mit Blick auf das Schreibziel und die intendierte Textsorte inhaltlich überarbeiten können sollten (D-EDK, 2016, S. 37). Dabei wird erwähnt, dass die Schüler\*innen positive Aspekte wie auch Unstimmigkeiten erkennen sollten: in Klasse 3–6 mithilfe von Kriterien, ab Klasse 7 im Austausch mit anderen. Feedback wird damit also curricular unterschiedlich modelliert.

Ein Blick bspw. in zwei Schweizer Deutsch-Lehrmittel für Sekundarstufe I (7.–9. Klasse) verweist auf eine andere Problemlage, die der Frage nach der Art der Vermittlung zugeordnet werden kann und die so für Deutsch-Lehrmittel in den drei Vollzentren Deutschland, Österreich und Schweiz generell zutreffen dürfte:

- *Satzweises Vorgehen*: In den Lehr-/Lernarrangements, die das inhaltliche Überarbeiten fokussieren, wird die Aufmerksamkeit der Schüler\*innen beim Überarbeiten hauptsächlich auf die Satzebene gelenkt, indem in erster Linie mit den grammatischen Satzproben gearbeitet wird (Sprachwelt Deutsch, 2012), wie dies auch bereits Baurmann und Ludwig (1996) vorschlugen. Wird der Blick auf das kommunikative Schreibziel gerichtet wie im Lehrmittel *Die Sprachstarken* (2015), wird dennoch satzweise vorgegangen, indem jeder Satz einzeln geprüft werden soll.<sup>3</sup> Das satzweise Vorgehen bzw. die Ausrichtung an

3 Das Lehrmittel orientiert sich dabei am Verfahren CDO (Compare – Diagnose – Operate) (vgl. MacArthur, 2012, S. 465).

den Satzproben ist auch in Umsetzungen mit Peer-Feedback-Verfahren dominant, wie sie in den beiden Lehrmitteln angeboten werden, so in der Schreibkonferenz oder der Textlupe – teilweise auch als „Schreibkarussell“ bezeichnet. Eine Aufmerksamkeitslenkung auf die Satzebene führt jedoch dazu, dass hauptsächlich auf der Wort- und Satzebene überarbeitet wird (MacArthur, 2012; McCutchen, Francis & Kerr, 1997). Globale Überarbeitungen auf der Textebene werden damit also kaum initiiert, sodass sich die Überarbeitung nicht oder nur in geringerem Maße auf die Textqualität auswirken dürfte.

- *Implizite Instruktion*: Eine explizite Vermittlung einer Überarbeitungsstrategie wird etwa im Kommentarband des Lehrmittels *Die Sprachstarken* für die 9. Klasse skizziert, dürfte aber – in Analogie zur formulierten Kritik in Applebee und Langer (2011) sowie Dignath und Veenman (2021) – eher selten von Lehrpersonen umgesetzt werden (vgl. auch Sturm, Lindauer & Sommer, 2023). Das Überarbeiten wird v. a. im Sinne eines ko-konstruktiven Zugangs angeleitet. Das gilt nicht zuletzt für die Schreibkonferenz oder die Textlupe, die über die alternativ-obligatorischen Deutsch-Lehrmittel hinaus bekannt sind und angeboten werden.

Für Verfahren, die eine Überarbeitungsstrategie explizit vermitteln oder die die Rolle der Leser\*innen stärker modellieren, lässt sich je nach Umsetzung ein teilweise sehr starker positiver Effekt auf den Lernzuwachs der Schüler\*innen nachweisen, gerade auch in Kombination (MacArthur, 2016; Rijlaarsdam et al., 2008). Beide Verfahren sind jedoch in Lehrmitteln kaum anzutreffen.

Zwei Studien konturieren dieses Bild nochmals:

Eine Studie zur schulischen Leistungsentwicklung u. a. in Deutsch, die von Angelone, Keller und Moser (2013) im Auftrag der Bildungsdirektion Kanton Zürich durchgeführt wurde, misst die Fähigkeit, *Texte* zu überarbeiten, ausschließlich auf der Wort- und Satzebene, indem die Schüler\*innen zu vorgegebenem Sprachmaterial bspw. passende Konjunktionen oder Zeitformen einsetzen sollen. Aufgaben dieses Typs werden durchschnittlich zu 63 % korrekt beantwortet. In der Leistungsentwicklung von Klasse 3 bis 9 lässt sich erwartungsgemäß ein Effekt der sozialen Herkunft feststellen, indem Schüler\*innen mit benachteiligter sozialer Herkunft in den leistungsschwachen Gruppen stark übervertreten sind (Angelone et al., 2013, S. 72ff.).

Eine Evaluation des Schwerpunkt-Programms *Schreiben auf allen Schulstufen*, das an Zürcher Schulen mit einem hohen Anteil an multilingualen Schüler\*innen durchgeführt wurde, deutet des Weiteren darauf hin, dass selbst Verfahren wie die Schreibkonferenz eher selten angewandt werden (Roos, 2017, S. 36): So gaben die befragten Lehrpersonen ( $N=2115$ ) an, Textlupe oder Schreibkonferenz als Peer-Feedback-Verfahren eher nicht oft zu verwenden; dagegen lassen sie die Schüler\*innen eher oft die von ihnen verfassten Texte vorlesen. Teilweise äußerten sie sich generell kritisch gegen Peer-Feedback:

In offenen Textantworten wird allerdings zum Peer-Feedback verschiedentlich angemerkt, dieses funktioniere nicht, weil die meisten Schülerinnen und Schüler sprachlich selber zu wenig stark seien, um anderen gezielte Rückmeldungen zu ihren Texten geben zu können. (Roos, 2017, S. 22)

Ein wichtiges Ziel des Schwerpunkt-Programms *Schreiben auf allen Schulstufen* war es u. a., mit schulinternen Weiterbildungen sowie ausführlichen Lehr-/Lernsettings das Repertoire der Lehrpersonen auszubauen (Sturm, 2023). In der Evaluation von Roos (2017) zeigte sich, dass die Lehrpersonen zwei Jahre nach Start des Programms vermehrt angaben, Peer-Feedback zu nutzen ( $d = .22$ ), während die Angaben zum summativen Beurteilen stabil blieben. Um das Repertoire auszubauen, wurde auf Verfahren gesetzt, die sich in der Literatur als besonders wirksam erwiesen haben.

## 4. Feedbackformen und instruktionale Prinzipien

Meta-Analysen wie Graham et al. (2011) oder Koster et al. (2015) können wie bereits erwähnt positive Effekte von (formativem) Feedback auf die Schreibleistung berichten. Um zu verstehen, was warum wie wirkt, lohnt sich ein Blick in die einzelnen Studien. Der folgende Überblick bezieht in erster Linie Interventionsstudien ein – auch solche, die in den Meta-Analysen (noch) nicht enthalten sind –, die das Überarbeiten kombiniert mit Feedback-Verfahren fokussieren.<sup>4</sup> Dabei liegt ein Schwerpunkt auf den instruktionalen Prinzipien, die die einzelnen Verfahren nutzen. Um diese besser herausarbeiten zu können, werden zudem Verfahren wie in Yarrow und Topping (2001) oder Knott (2020), die Überarbeiten und Feedback als integralen Bestandteil des gesamten Textproduktionsprozesses modellieren, nicht berücksichtigt.

### 4.1 Schriftliches Feedback

Verschiedene Studien haben *kriterienorientiertes schriftliches Feedback* überprüft: Sie zeigen auf, dass sich eine Verbesserung erzielen lässt, wenn die Schüler\*innen einen genrespezifischen Kriterienkatalog – passend zur Schreibaufgabe – zur Verfügung haben, mit welchem sie ihre Texte überarbeiten. Zu nennen ist hier exemplarisch die Studie von Philippakos und MacArthur (2016). Sie teilten Schüler\*innen der Klasse 4/5 in zwei Gruppen ein:

- a) Die Schüler\*innen evaluieren an drei hintereinander folgenden Tagen jeweils zwei argumentative Texte von Peers mithilfe eines Kriterienkatalogs und formulieren ein schriftliches Feedback.
- b) Analog wie Gruppe a), aber ohne Formulierung eines Feedbacks.

<sup>4</sup> Rekonstruierende Studien sowie Fallstudien werden deshalb weitgehend ausgeblendet (vgl. dazu etwa den Sammelband Hüttis-Graf & Jantzen, 2012).



Für die Gruppe a) zeigte sich gegenüber der Gruppe b) ein mehrfacher signifikanter Vorsprung, so etwa in Bezug auf die Qualität der durchgeführten Revisio-  
nen sowie die Qualität der eigenen Texte, die nach dem Evaluieren und Formu-  
lieren von Feedback verfasst wurden. Allerdings zeigte sich gleichzeitig, dass sich  
der Vorsprung bereits zwei Wochen später wieder verringerte, was so interpretiert  
werden kann, dass es mehrfache Übungssituationen zur Festigung des Gelernten  
braucht.

Angemerkt sei zudem, dass Philippakos und MacArthur (2016) einen Krite-  
rienkatalog verwenden, der sich an der prototypischen Struktur des Zieltextes  
orientiert (vgl. Tabelle 1). Becker-Mrotzek und Böttcher (2012) sehen zwar eben-  
falls textsortenspezifische Kriterienraster vor, orientieren sich aber an einem  
Basiskatalog, der die Kriterien nach Sprachrichtigkeit, Sprachangemessenheit,  
Inhalt, Aufbau und Prozess ordnet. In beiden Fällen sind zur Evaluation jeweils  
drei Ausprägungen vorgesehen, die Philippakos, MacArthur und Coker (2015) zu-  
sätzlich mit „nicht vorhanden“, „könnte besser sein“ und „großartig!“ beschriften.

Tabelle 1: Kriterienkatalog argumentative Texte (Philippakos et al., 2015; Übersetzung  
A.S.)

|            |                                                                                                                                                                                                                                 |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Einleitung | <i>Thema:</i> Was ist das Thema und weshalb sollte sich der Leser, die Leserin dafür interessieren?<br><i>Meinung:</i> Ist klar, welche Meinung der Autor, die Autorin vertritt?                                                |
| Mittelteil | <i>Argument 1:</i> Wird das erste Argument mit der Meinung verknüpft? Ist es für die Lesenden klar und überzeugend?                                                                                                             |
|            | <i>Evidenz:</i> Wird das Argument ausreichend begründet und mit Beispielen gestützt?                                                                                                                                            |
|            | <i>Argument 2:</i> ...                                                                                                                                                                                                          |
| Schluss    | <i>Meinung:</i> Wiederholt und reformuliert der Autor, die Autorin die eigene Mei-<br>nung?<br><i>Nachdenken:</i> Gibt der Autor, die Autorin den Lesenden eine Botschaft mit, die sie<br>zum Nachdenken über das Thema anregt? |
| Anderes    | [Titel, sprachlicher Stil, sprachliche Angemessenheit]                                                                                                                                                                          |

Die Studie von Philippakos und MacArthur (2016) kann nicht nur als Beleg für  
die Effektivität des kriterienorientierten Feedback in Bezug auf die Schreibkompe-  
tenz gesehen, sondern sie kann auch so gelesen werden, dass die Qualität des Kri-  
terienkatalogs eine wichtige Rolle spielt.

## 4.2 Schreibgespräche

Von kriterienbasierten Verfahren sind Förderansätze zu unterscheiden, die die *Schreibumgebung dialogisch gestalten*: Dazu zählt in erster Linie die Schreibkonferenz, die in den Deutsch-Lehrmitteln einen festen Platz einnimmt und die vor allem auf Sprachhandlungen wie Vorlesen, Zuhören, Sprechen oder Aufschreiben setzt, „die im Dialog spontan, situationsverschränkt, affektiv etc. geäußert werden“ (Reichardt, Kruse & Lipowsky, 2014, S. 69). Die Entwürfe werden dabei in Gruppen u. a. unter inhaltlichen, stilistischen und orthografischen Aspekten besprochen. Die Schüler\*innen (3. Klasse) der Gruppe Schreibkonferenz überarbeiteten ihre Texte anschließend effektiver als jene, die mit der Textlupe arbeiteten (vgl. Kap. 3.3). Der Effekt ist mit  $\eta^2 = .05$  jedoch gering.

Held (2006) oder Allal (2018) leiteten in ihren Studien zu Schreibgesprächen in Dyaden an und konnten nicht zuletzt im Rahmen einer qualitativen Analyse Vorteile gegenüber einer Kontrollgruppe herausarbeiten.

Angemerkt sei, dass die Autorenrunde von Leßmann (2021) als Variante der Schreibkonferenz betrachtet werden kann, die aber wöchentlich und im Zusammenhang mit freiem Schreiben praktiziert werden soll. Zu diesem Ansatz liegen Fallstudien, jedoch keine Interventionsstudien (mit Kontrollgruppe) vor.

## 4.3 Leser\*innen beobachten

Dass die Schreibkonferenz wie zuvor ausgeführt zu besseren Ergebnissen als die Textlupe führt, könnte daran liegen, dass die Schreibkonferenz im direkten Dialog stattfindet, sodass die Lesenden beobachtbar werden. Eine weitere Gruppe von Förderansätzen stellt denn auch das *Beobachten von Leser\*innen* ins Zentrum. Im Unterschied zur Schreibkonferenz erfolgt das Beobachten gewissermaßen nicht beiläufig, sondern wird in spezifischer Weise herausgefordert. In der Regel werden dazu Textsorten eingesetzt, die zu Handlungen auffordern oder anleiten, vor allem instruktive Texte.

Feedback mit Beobachten von Leser\*innen lässt sich grob in a) synchrone und b) asynchrone Verfahren gliedern.

### a) synchrone Verfahren

Synchrone Verfahren lassen sich des Weiteren danach unterscheiden, ob sie Tandems oder kleine Gruppen bilden:

*Tandems*: Ein Schreiber beobachtet eine Leserin, die eine Kopie des Textes vor sich hat. Die Leserin denkt laut beim Lesen des instruktiven Textes und führt die Anleitung dabei direkt aus. Der Schreiber markiert sich in seinem Text laufend Stellen, die Probleme aufwerfen, und notiert sich evtl. bereits mögliche Revisionen.

Vertreter dieses Ansatzes sind Couzijn und Rijlaarsdam (2005) oder López, Torrance, Rijlaarsdam und Fidalgo (2021). Die Studien belegen hohe Effektstärken

auf die Textqualität (ab  $d = .80$  bis  $d = 1.48$ ) und mittlere bis hohe Effektstärken auf einzelne Aspekte wie Textstruktur oder substanzielle Revisionen (ab  $d = .70$ ).

*Gruppe:* Es werden zwei Gruppen gebildet: die Redaktion (=Leser\*innen) und eine Art Zuschauertribüne – Amir et al. (2021) bezeichnen diese Gruppe als Forscher\*innen –, die die Leser\*innen beobachtet und sich notiert, welche Kriterien die Leser\*innen verwenden (Rijlaarsdam et al., 2008). Es handelt sich dabei um eine Variante der sog. Fishbowl. Die Redaktion hat die Aufgabe, einen Text aus mehreren Texten auszuwählen und genau zu begründen, was sie an einem Text (nicht) überzeugt. Wenn die Redaktion ihre Aufgabe beendet hat, berichtet die Zuschauertribüne den Leser\*innen, welche Kriterien sie beobachtet hat.

Zentral dabei ist, dass die Schüler\*innen nicht vorab Kriterien erhalten, sondern dass sie selbst mit Bezug zur Schreibfunktion Kriterien entwickeln. Im Verlauf einer solchen Redaktionssitzung erweitern die Schüler\*innen laufend ihre Kriterien, indem sie mit jedem weiteren Text alternative Lösungen erhalten und so die Texte bzw. verschiedenen Lösungen zu vergleichen beginnen. In der Regel entscheiden sich die Schüler\*innen nicht vorbehaltlos für einen Text, sondern sie wählen einen Text aus und formulieren Überarbeitungsvorschläge, und zwar mit Blick auf andere Texte, die bestimmte Elemente aus ihrer Sicht besonders gut umgesetzt haben.

Angemerkt sei, dass sich solche synchronen Verfahren von einer Schreibkonferenz unterscheiden, indem Schreibende und Lesende nicht direkt interagieren, sondern die Schreibenden und Lesenden getrennt werden, sodass die Reaktion der Leser\*innen besser beobachtbar ist. Des Weiteren muss ein Text aus einer Gruppe von Texten ausgewählt werden, was einen Vergleich der verschiedenen „Lösungen“ bedingt, was bspw. in der Schreibkonferenz nicht angelegt ist. Es ist in solchen Redaktionssitzungen denn auch beobachtbar, dass die Redaktionsmitglieder im Verlauf ihrer Diskussion immer wieder auf Texte zurückgreifen, die bereits besprochen wurden, und auf diese Weise zudem sehr textnah argumentieren.

Moore und MacArthur (2012) setzen zwar ein ähnliches Verfahren ein, lassen aber nicht einen Text aus mehreren Texten auswählen, sondern die Schüler\*innen in der Rolle als Redaktionsmitglied sollen diskutieren, inwiefern der einzelne Text die Lehrperson überzeugen könnte. Hierzu erhalten sie Leitfragen, müssen also nur bedingt selbst Kriterien entwickeln. Ein Vergleich der verschiedenen Lösungen fällt bei dieser Version ebenfalls weg.

Für die Zuschauertribüne zeigt sich Rijlaarsdam et al. (2008, S. 63) zufolge mit  $d = 1.30$  ein grosser Effekt auf die Textqualität, während sich für die Redaktion nur geringere Effekte nachweisen lassen ( $d = .30$ ). Eine Replikationsstudie von Amir et al. (2021, S. 19) mit einer größeren Stichprobe zeigt ähnliche Effekte ( $d = .95$  vs.  $.29$ ). Für die abgewandelte Form von Moore und MacArthur (2012, S. 1461) zeigt die Redaktionsgruppe im Vergleich zur Kontrollgruppe bei der ersten Aufgabe bessere Schreibleistungen ( $d = .62$ ), während sich die Zuschauertribüne nicht signifikant von der Redaktions- und Kontrollgruppe unterscheidet. Bei einer zweiten Aufgabe dagegen, die Moore und MacArthur (2012, S. 1463) als Transfer-

aufgabe bezeichnen. lassen sich zwischen allen drei Gruppen keine Unterschiede mehr nachweisen.

Die Unterschiede zwischen den beiden Versionen erklären Amir et al. (2021, S. 19f.) hauptsächlich damit, dass sie selbst die Intervention in 7. Klassen, Moore und MacArthur (2012) dagegen in 5. Klassen durchführten. Denkbar wäre aber auch, dass der Vergleich der alternativen Lösungen sowie die selbstständige und laufende Erarbeitung von Kriterien entscheidend zur Wirksamkeit des Verfahrens beiträgt und dass die Orientierung daran, was die Lehrperson überzeugen könnte, zu besonderen Schwierigkeiten geführt hat.

### *b) asynchrone Verfahren*

Mehrere Studien wählen ein Vorgehen, indem sie Schüler\*innen in Einzelarbeit ein produktorientiertes Feedback geben lassen: Die Leser\*innen lesen bspw. einen instruktiven Text und führen anschließend die Anleitung gemäß Text durch. Die Schreiber\*innen beobachten also die Leser\*innen nur indirekt, indem sie das Produkt der Leser\*innen – eine Zeichnung, ein Tangrambild o. Ä. – erhalten. Die Schreiber\*innen analysieren dann selbst mithilfe der Lösung der Leser\*innen, welche Stellen ihres Textes Probleme aufgeworfen haben.

Exemplarische Vertreter\*innen dieser Richtung sind u. a. Anskit (2019), Gan und Hattie (2014) oder Traxler und Gernsbacher (1993). Insgesamt weisen die einzelnen Studien sehr unterschiedliche Befunde auf (von  $d = .28$  bis  $d = 1.28$ ). Da das direkte Beobachten der Lesenden wegfällt, dürften hier andere instruktionale Merkmale entscheidender sein, sei es die Einbettung der Schreibaufgabe in einen Handlungsrahmen, das wiederholte Zuordnen verschiedener Lösungen zu einem Text oder in neueren Arbeiten wie bei Anskit (2019) die Profilierung der Schreibarrangements (inkl. Textprozeduren).

Einen Spezialfall des asynchronen Zugangs stellt die Studie von Lumbelli, Paoletti und Frausin (1999) dar: Die Schreiber\*innen lesen zuerst ihren eigenen Text laut und hören danach eine Audioaufnahme, in der ein\*e (erwachsene) Leser\*in zu einem ähnlichen Text laut denkt und dabei auf fehlende Informationen oder Lücken im Text aufmerksam macht. Anschließend übertragen die Schreiber\*innen das Gehörte auf ihren eigenen Text. Dieses Feedback-Verfahren zeigt einen großen positiven Effekt von  $d = .83$  auf die Schreibleistung der Lernenden.

Die Textlupe kann zwar auch als ein asynchrones Verfahren in Einzelarbeit bezeichnet werden: Während die hier berichteten asynchronen Verfahren die Wirkung des Textes auf globaler Ebene fokussieren, operiert die Textlupe dagegen eher auf der lokalen Ebene und vor allem auch sehr unspezifisch, indem u. a. unabhängig von der intendierten Textsorte allgemeine Aspekte wie „Das gefällt mir an deinem Text“ oder „Dazu habe ich Fragen“ vorgegeben werden. Reichardt et al. (2014, S. 77) berichten für die Textlupe mit  $\eta^2 = -.03$  denn auch einen geringen negativen Effekt auf die Textqualität.