



Beurteilung argumentativer Schüler*innentexte durch Geographielehrkräfte

Saskia Steingrübl, Alexandra Budke
(Institut für Geografiedidaktik, Universität zu Köln)

Abstract:

Textbeurteilungen sowie implizierte Fehlerkorrekturen sind erste Schritte zur Diagnose von Schüler*innenkompetenzen, anschließender Rückmeldung und Förderung. Insbesondere das argumentative Schreiben stellt im Geographieunterricht eine große Relevanz, jedoch gleichzeitig eine Herausforderung dar – Geographielehrkräfte müssen hierbei vielerlei Kriterien bei der Beurteilung berücksichtigen. Im Zuge einer explorativen, qualitativen Studie anhand von leitfadengestützten Interviews und Videoanalysen von digitalen Textbeurteilungen wird untersucht, wie Geographielehrkräfte bei der Korrektur argumentativer Schüler*innentexte vorgehen, welche Korrekturmethode und Beurteilungskriterien sie auswählen und welche Einstellungen zur Funktion der Textbeurteilung sie haben. Es wird deutlich, dass sprachliche Kriterien bei der Korrektur im Vordergrund stehen, die Korrekturmethode stark variieren, lehrkraftabhängig und nicht standardisiert sind.

Text assessment and implicit error correction are the first steps in diagnosing students' competence, subsequent assessment, feedback and support. In particular, argumentative writing is very relevant in geography education, but at the same time it is challenging - geography teachers have to consider many criteria when assessing it. An exploratory, qualitative study based on guided interviews and video analysis of digital text assessments examines how geography teachers go about correcting students' argumentative writing, what correction methods and assessment criteria they choose, and what attitudes they have towards the function of text assessment. It becomes clear that linguistic criteria are at the forefront of correction, that correction methods vary widely, depend on the teacher and are not standardised.

Keywords: Textbeurteilung, argumentative Texte, Beurteilungskriterien, Geographielehrkräfte, Lehrkraftkompetenzen, Textkorrektur

Empfohlene Zitierweise:

Steingrübl, S., & A. Budke (2023): Beurteilung argumentativer Schüler*innentexte durch Geographielehrkräfte. zisch: zeitschrift für interdisziplinäre schreibforschung, 9, 69-100. DOI: <https://doi.org/10.48646/zisch.230904>



Lizenziert unter der CC BY-ND 4.0 International Lizenz.

Dieses Werk ist lizenziert unter einer [Creative Commons Namensnennung - Keine Bearbeitungen 4.0 International Lizenz](http://creativecommons.org/licenses/by-nd/4.0/) zugänglich. Um eine Kopie dieser Lizenz einzusehen, konsultieren Sie <http://creativecommons.org/licenses/by-nd/4.0/> oder wenden Sie sich brieflich an Creative Commons, Postfach 1866, Mountain View, California, 94042, USA.

Beurteilung argumentativer Schüler*innentexte durch Geographielehrkräfte

Saskia Steingrübl, Alexandra Budke
(Institut für Geografiedidaktik, Universität zu Köln)

Einleitung

Rotstift, Wegstreichen, Dranschreiben – Textbeurteilungen gehören zum Alltag der meisten Geographielehrkräfte. Schüler*innen produzieren Texte, Lehrkräfte sammeln diese teilweise ein, korrigieren und geben sie zurück. Es werden dabei unterschiedliche Textsorten von den Schüler*innen geschrieben, u.a. Argumentative. Ein argumentativer Text ist dabei eine schriftliche Darstellung von Thesen, die durch Geltungsbeziehungen mit Daten verknüpft werden, um die Adressat*innen zu überzeugen (Toulmin, 1996; Kopperschmidt, 2000). Das Ziel beim Schreiben argumentativer Texte im Geographieunterricht ist unter anderem die Förderung systemischen Verständnisses für komplexe Sachverhalte, der fachlich begründeten Urteilsbildung, der Vermittlung von Argumentationskompetenzen und sozialen Kompetenzen, der Problemlösungsfähigkeit sowie die Vorbereitung auf die Teilhabe an gesellschaftlichen Debatten und demokratischen Prozessen (Budke, 2020; Budke & Meyer, 2015 u.v.m). Dazu müssen Schüler*innen oftmals ökologische, ökonomische, soziale, räumliche und zeitliche Beurteilungskriterien in Verbindung bringen, abwägen und in ihrer Bedeutung für die Fragestellung beurteilen und kritisch Stellung nehmen. Studien zeigen, dass Schüler*innenkompetenzen in diesem Bereich ausbaufähig sind und daher im Unterricht gefördert werden sollten (Zohar & Nemet, 2002; Sandoval & Millwood, 2005; Sampson & Clark, 2007; Von Aufschnaiter et. al., 2008; Budke & Uhlenwinkel, 2011; IQB, 2014; Pohl, 2014; Budke & Kuckuck, 2017). Die Korrektur von argumentativen Texten kann zum einen dazu dienen, Entwicklungsstände zu diagnostizieren und zielgenau Defizite zu fördern und zum anderen die Leistung zu beurteilen und Rückmeldung zu geben. An dieser Stelle stellt sich die Frage, welche Kriterien Lehrkräfte bei der Beurteilung von argumentativen Schüler*innentexten beachten und wie sich der Beurteilungsvorgang gestaltet. Bislang gibt es in der Geographiedidaktik lediglich Studien, welche die Schüler*innenkompetenzen bezüglich argumentativer Texte untersucht haben (Budke et. al., 2010; Budke & Uhlenwinkel, 2011; Budke & Kuckuck, 2017) und eine Untersuchung, die sich auf die Beurteilung von Abiturklausuren bezieht (Weiss et. al., 2013). In weiteren Fachdidaktiken zeigen Studien, welche Kriterien Lehrkräfte bei der Beurteilung von Texten (Busch & Ralle, 2013; Rüede & Weber, 2009; Feser & Höttecke, 2016), welche jedoch nicht argumentativ sind, anwenden. In den Sprachdidaktiken spielt die Thematik der Lehrkraftkompetenzen bezüglich der Textbeurteilung eine größere Rolle. Dennoch gelten für den Fachunterricht teils andere Beurteilungskriterien, welche die Fachinhalte in den Vordergrund stellen, was sowohl die Bildungsstandards und Kernlehrpläne als auch die daraus resultierenden Bewertungskriterien für die Abituraufgaben zeigen (KMK, 2005; DGfG, 2020). Somit handelt es sich bei der nachfolgenden Untersuchung der Lehrkraftkompetenzen bezüglich der

Textbeurteilung argumentativer Schüler*innentexte im Geographieunterricht um ein Desiderat. In der Studie werden somit folgende Forschungsfragen untersucht:

- Wie gehen Geographielehrkräfte bei der Beurteilung von argumentativen Schüler*innentexten vor?
- Welche Kriterien nutzen Geographielehrkräfte zur Beurteilung argumentativer Schüler*innentexte?
- Welche Methoden verwenden Geographielehrkräfte zur Beurteilung argumentativer Schüler*innentexte?

Zehn Geographielehrkräfte wurden im Rahmen qualitativer Leitfadeninterviews zunächst zu ihrer Selbsteinschätzung und ihrem Unterrichtshabitus bezüglich der Beurteilung argumentativer Schüler*innentexte befragt. Anschließend wurden sie gebeten, einen argumentativen Schülertext zu korrigieren. Die Gedanken beim Beurteilungsprozess der Lehrkräfte wurden durch Laut-Denken parallel zur Videographie der Textkorrektur erfasst.

Im Folgenden werden zunächst die Herausforderungen beim argumentativen Schreiben dargestellt sowie Studienergebnisse zum Thema präsentiert. Daran anknüpfend erfolgt eine definitorische Abgrenzung der Begriffe ‚Korrektur‘, ‚Textbeurteilung‘ und ‚Textfeedback‘. Anschließend werden Methoden und Grundsätze der Textbeurteilung dargelegt und der aktuelle Forschungsstand zu Schüler*innentextbeurteilungen im Fachunterricht aufgezeigt sowie der aktuelle Diskurs und Forschungsstand der Textbeurteilungskompetenzen von Lehrkräften im Fachunterricht beschrieben. Daraus resultierend wird die Forschungsmethodik erläutert und es werden literaturbasiert Kriterien für die Textbeurteilung argumentativer Schüler*innentexte entwickelt. Schlussendlich werden die Ergebnisse der Studie präsentiert und diskutiert sowie ein Ausblick für die (Hoch-)Schullehre und Forschung gegeben.

Theoretische Hintergründe

Schriftliche Argumentation als Herausforderung im Unterricht

Schriftliche Argumentationen sind aus dem Fachunterricht nicht wegzudenken. Speziell im Geographieunterricht sind sie ein Mittel, um verschiedene Akteur*innenpositionen, Raumnutzungskonflikte und Umwelt- und Gesellschaftsprobleme anhand von diversen kontinuierlichen und diskontinuierlichen Materialien auszuhandeln. Dabei werden vielfältige Konflikte auf unterschiedlichen räumlichen und zeitlichen Maßstabsebenen behandelt, um kreative Lösungsansätze zu finden, darüber zu informieren und Rezipient*innen von der eigenen Meinung zu überzeugen. Diese kommunikative Aushandlung von Problemen spielt eine wichtige Rolle bei der gesellschaftlichen Teilhabe (Budke & Meyer, 2015). Zudem ist Argumentation ein Bestandteil der Wissenschaft, welcher dazu dienen kann, Sachverhalte zu erklären, zu diskutieren und weiterzudenken, um zu Lösungsansätzen von aktuellen oder zukünftigen Problematiken zu gelangen (Erduran & Jiménez-Aleixandre, 2007). Dies gilt dabei für alle Fachdidaktiken, denn das schriftliche Argumentieren dient insbesondere dazu, das Verständnis fachlicher Inhalte zu stärken (Budke & Meyer, 2015). Dabei ist das schriftliche Argumentieren eine äußerst komplexe Handlung, bei der viele Teilkompetenzen zum Einsatz kommen: Schüler*innen

müssen dabei beispielsweise über das fachsprachliche Vokabular, argumentative Handlungsschemata und Wissen über typische Textstrukturen verfügen. Im Bereich der methodischen Kompetenzen müssen sie fachbezogene Argumentationsmuster sowie fachtypische Gütekriterien anwenden können. Dabei ist die Materialauswertung ein wichtiger Bestandteil, um fachlich gültige und passende Belege für die eigenen Argumente zu nutzen. Darüber hinaus müssen Schüler*innen dabei Vorwissen über fachinhaltliche Belange zu Akteur*innen, Positionen und Fakten und „Wissen über fachliche Relevanz der Fragestellung, fachlichen Eignung von Belegen, fachliche Gültigkeit von Geltungsbeziehungen“ (Budke, 2021) verfügen. Während sprachliche Register überwiegend fächerübergreifend ähnlich und relevant sind, so sind die methodischen und fachinhaltlichen Anforderungen der Fächer verschieden und unterschiedlich gewichtet. So spielt beispielsweise das kartenbasierte Argumentieren und die damit einhergehenden fachlichen Kompetenzen im Geographieunterricht eine relevantere Rolle als beispielsweise in der Deutschdidaktik. Im Folgenden werden die Herausforderungen von Schüler*innen beim schriftlichen Argumentieren dargestellt, welche sich aus der großen Anzahl an benötigten Teilkompetenzen ergeben.

Studien zum schriftlichen Argumentieren in den Fachdidaktiken zeigen, dass Schüler*innen Argumente teilweise formulieren können, jedoch Schwierigkeiten mit dem multiperspektivischen Argumentieren, dem Materialbezug (Hindmarsh & Budke, 2023) und bei der Einbindung relevanter Informationen aus dem Fachunterricht haben (Budke & Uhlenwinkel, 2011; Sandoval & Millwood, 2005; Kelly & Bazerman, 2003). Diese werden selten mit eigenen Ideen verknüpft (Sandoval & Millwood, 2005) und die Argumente unzureichend durch Belege gestützt (ebd. ; Zohar & Nemet, 2002). Des Weiteren wurden Probleme mit dem Bilden eines ‚roten Fadens‘ innerhalb ihres Textes sowie der sinnvollen Verknüpfung von Inhalten festgestellt (Sampson & Clark 2008). Erkenntnisse aus der Deutschdidaktik bezüglich der Schüler*innenkompetenzen beim argumentativen Schreiben zeigen, dass Schüler*innen Schwierigkeiten beim Verfassen dieser Textsorte haben (Spiegel 2006; Petersen 2013; Stimming 2019): beispielsweise bei der Strukturierung der Argumentation (Becker-Mrotzeck, 2017).

Über die Referenzgruppe der Schüler*innen hinaus, gibt es in der Geographiedidaktik ebenso Studien zur schriftlichen Argumentationskompetenz, welche die Kompetenzen angehender Geographielehrkräfte untersuchten. Diese zeigten, dass die Studierenden zwar mehr geographisch relevante Argumente in ihren Texten verarbeiteten als die Schüler*innen (Budke & Uhlenwinkel, 2011), dennoch Schwierigkeiten bei der Einbindung von Gegenargumenten, Formulierung von Quellenbezügen und -verweisen, der genaueren Beschreibung der Bedingungen, unter denen die Argumente gelten und der Integration von Raumbezügen sowie einer generellen multiperspektivischen Darstellung des Sachverhaltes haben (Budke & Kuckuck, 2017). Des Weiteren gelingt es ihnen nicht immer logische Verbindungen zwischen Behauptungen und Belegen herzustellen und Argumente vollständig darzustellen (Budke et. al., 2010).

Aus dem Forschungsstand geht hervor, dass bereits einige Studien zu Argumentationskompetenzen vorliegen. Daraus ergibt sich der Bedarf an Förderung. Diese sollte an der Textbeurteilung und Kompetenzdiagnose der Lehrkräfte anknüpfen. Dazu benötigen Lehrkräfte selbst unterschiedliche Kompetenzen: eigene Argumentationskompetenzen, Textbeurteilungskompetenzen und

Feedbackkompetenzen bezüglich argumentativer Texte, um Beurteilungskriterien ableiten zu können. Diesbezüglich liegen speziell für schriftliche Argumentationen im Geographieunterricht keine Studien vor. Die nachfolgende Studie wird hierbei auf die Textkorrektur und -beurteilung durch Geographielehrkräfte fokussieren. Wie die Begrifflichkeiten dabei unterschieden werden, zeigt das nächste Kapitel.

Lehrkrafthandlungen im Überblick: Korrektur, Beurteilung, Diagnose & Feedback

Korrektur, Textbeurteilung, Diagnose, Feedback – Zunächst scheint es sinnvoll bei der Vielzahl von Begrifflichkeiten einen definitorischen Überblick zu geben, um die Unterschiede und damit auch die Handlungsfelder von Lehrkräften aufzuzeigen. Zur Übersicht soll Abbildung 1 die einzelnen Lehrkraft- und Schüler*innenhandlungen verdeutlichen, abgrenzen und in eine Reihenfolge bringen.

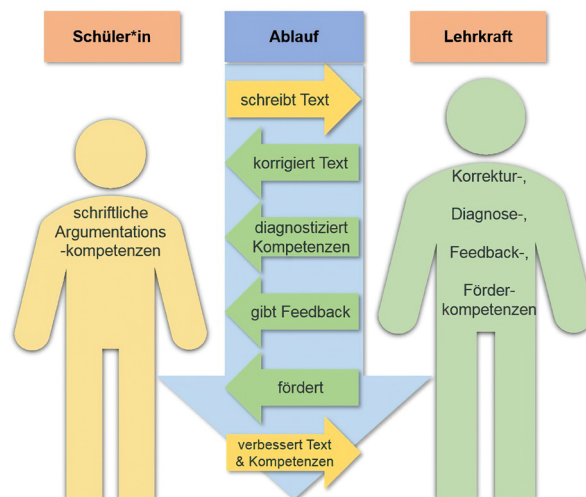


Abbildung 1 Relevanz und Zusammenhang der Schüler*innen- & Lehrkraftkompetenzen
(Quelle: eigene Darstellung)

Abbildung 1 zeigt die Zusammenhänge zwischen den Schüler*innenkompetenzen und ihrem Handeln beim Schreiben eines argumentativen Textes im Verhältnis zu den Lehrkraftkompetenzen im Umgang mit diesen Texten. Es handelt sich dabei um eine idealtypische Darstellung, welche die im Nachfolgenden zu definierenden Begriffe der Korrektur, Textbeurteilung, Diagnose und Feedback aufgreift.

Korrektur von Texten setzt dabei oftmals dort an, wo Fehler vorhanden sind. Dietmar Rösler (2012:151) versteht unter Fehlern „unbewusste Abweichungen der Norm“. Laut Duden bedeutet der Begriff der Korrektur so viel wie „Verbesserung; Berichtigung; Richtigstellung“ (Dudenredaktion, o.J.). Dabei werden Fehler im Text beseitigt und/oder verbessert und somit auch mehr die sprachliche Ebene fokussiert, die weniger konkrete Richtlinien und damit eine weniger komplexe Korrektur benötigt. Durch Textkorrekturen können die Problembereiche der Schüler*innen erkannt (Ferris, 2006), die Lernentwicklung gefördert (Ellis, 2009) und die Erkennung von Fehlermustern gestärkt (Bitchener &

Knoch, 2008) werden. Beurteilung definiert sich nach Müller und Ditton (2014:21) als „Bewertung einer Person, eines Verhaltens oder einer gezeigten Leistung auf der Basis von Beobachtungen“. Die Textbeurteilung ist nicht allein positiver konnotiert, sondern grenzt sich von der reinen Korrektur ab, indem sie keine Richtigstellung impliziert, jedoch eine Bewertung. Dazu muss sie zusätzlich die formelle und inhaltliche Beurteilung der beschriebenen Sachverhalte fokussieren. Eine Funktion der Textbeurteilung ist die Leistungsmessung durch die Lehrkräfte. Hierbei unterscheiden Bartosch & Mayr (2009) zwischen Leistungsmessung und Leistungsbewertung. So wird die Leistungsmessung als Betrachtung des „Ist-Zustandes“ anhand unterschiedlicher Kriterien bzw. geeigneter Messinstrumente verstanden. Bei der Leistungsbewertung werden die Erwartungen an den Soll-Zustand mit dem Ist-Zustand der Schüler*innenkompetenzen verglichen und ausgewertet. Lehrkräfte müssen somit für die Beurteilung von Schüler*innenleistungen zunächst Kriterien erarbeiten, die den Grad der Richtigkeit der Textprodukte der Schüler*innen bestimmt (Wieser, 2014) – dabei wächst die Komplexität der Faktoren mit dem Umfang und der Schwierigkeit der von den Schüler*innen zu bewältigenden Leistungen.

Der Begriff der Diagnose wird von Rüede und Weber (2009:819) als „das Sammeln und Werten von Informationen über einen bestimmten Sachverhalt zur Gewinnung eines Gesamtbildes“ bezeichnet. Diagnostik, welche in der Schule durchgeführt wird, dient immer der Optimierung von Lernprozessen und bildet den Ausgangspunkt für zu treffende Fördermaßnahmen und Binnendifferenzierung. Diagnose soll den Lehrkräften ermöglichen, Informationen über den Kompetenzstand der unterrichteten Schüler*innen zu geben, um dahingehend Entscheidungen über die Unterrichtspraxis zu treffen, sachgerechte Noten zu geben, Eltern und Kolleg*innen über die Leistungen informieren zu können und Empfehlungen bezüglich einer weiteren Förderung auszusprechen (Van Ophuysen & Lintorf, 2013). Helmke (2012) betont indes, dass Diagnose insbesondere der Informationsbeschaffung dient, mit dem Ziel, personengerechte Entscheidungen zur individuellen Förderung sowie Anpassungen des Unterrichtsangebotes zu treffen. Dabei ist Diagnose zunächst wertfrei.

Feedback oder Rückmeldung hingegen ist individuell, sollte verständlich formuliert sein und die Lernenden dazu anregen, ihren Lernprozess zunehmend selbst zu steuern (Rae & Cochrane, 2008). Rückmeldung zu Texten sollte bestenfalls positive Aspekte beinhalten, als auch über Fehler und Problembereiche informieren. Diese werden aus pädagogischer und sprachwissenschaftlicher Sicht als relevanter Teil des Lernens angesehen, wenn sie als Feedback über die Schüler*innenleistung verarbeitet werden (Hattie, 2012).

In der nachfolgenden Studie wird der Fokus auf die Textbeurteilungspraktiken von Geographielehrkräften hinsichtlich eines argumentativen Schülertextes gelegt. Dabei wird Textbeurteilung und Leistungsbeurteilung synonym verwendet und die curriculare Funktion, welche das Ziel der bestmöglichen Förderung für Schüler*innen hat und der Optimierung des Unterrichts durch die Lehrkraft gilt (Jürgens, 2010), in den Vordergrund gestellt.

Methoden und Kriterien der Textbeurteilung

Entgegen der ganzheitlich-intuitiven Verfahren zur Textbeurteilung, geschieht diese bestenfalls kriterienorientiert und sichert Transparenz, Akkuratheit und Differenzierung der Fehlerarten. Heid (2015:35) beschreibt hierzu: „Es gibt [...] keine Fehler „an sich“, sondern nur Sachverhalte, die mit

Bezug auf ein entscheidungsabhängiges Bewertungskriterium (eine Norm) von einem Subjekt als fehlerhaft beurteilt werden“. Dies zeigt, dass Beurteilungskriterien eine Beschreibung der Fehlerquellen bieten und dadurch eine präzisere Kompetenzdiagnose und der Rückmeldung ermöglichen. Ebenso wird die gemeinsame Erarbeitung von Kriterienkatalogen mit Schüler*innen im Unterricht als sinnvoll angesehen, um auf Vorwissen zurückzugreifen (Becker-Mrotzeck, 2020) und Leistungstransparenz zu sichern.

Becker-Mrotzeck (2020) beschreibt, dass förderorientiertes Korrigieren und Kommentieren im Sinne der Lehrkraft sein sollte. Diese sollte Schreiben als Prozess sehen, Kommentare adressatengerecht formulieren, Lernstände diagnostizieren können und über Selbstreflexion bezüglich der von ihnen gestellten Anforderungen der Schüler*innen gegenüber verfügen. Kleppin (2007) benennt als Korrekturverfahren „die einfache Fehlermarkierung“, „die Fehlermarkierung mit Korrekturzeichen“ (ebd.) sowie „die Berichtigung durch den Lehrer (sic! *in)“. Becker-Mrotzeck (2020) ergänzt und empfiehlt die Verwendung von Randkommentaren oder Kommentaren unter dem Text, damit der Eingriff in den Schüler*innentext vermieden wird und ein Dialog entstehen kann. Dazu legt er einige Maxime fest, welche in den Kommentaren berücksichtigt werden sollten: Diese umfassen beispielsweise die direkte Ansprache des Schülers/ der Schülerin, positive Anmerkungen, Begründungen und Verständnisfragen, Verständlichkeit der Kommentare und Lernanregungen (ebd.). Textbeurteilung sollte dabei zur Lernberatung eingesetzt werden und motivierend sein (Beer, 2006). Jürgens (2010) beschreibt darüber hinaus, dass Leistungsbewertungen dem Zweck der Selbstkontrolle dient und die Lernenden über ihren Leistungsstand aufklären soll.

Aktueller Forschungsstand zu Textbeurteilungen durch Lehrkräfte im Fachunterricht

Aus fachdidaktischen Studien in der Chemie, Physik, Mathematik und Geographie konnten einige Erkenntnisse bezüglich der Beurteilungskriterien der (angehenden) Fachlehrkräfte zu schriftlichen Schüler*innenprodukten gesammelt werden. Dabei wurde für den Chemieunterricht festgestellt, dass bei der Diagnostik von fachsprachlichen Kompetenzen insbesondere das Einbinden von Fachbegriffen für die Lehrkräfte im Mittelpunkt steht (Busch & Ralle, 2013). Rüede und Weber (2009) konnten für den Physikunterricht feststellen, dass angehende Lehrkräfte insbesondere auf formale, inhaltliche und Hintergrundaspekte aus beschreibender und wertender Perspektive achten. Die Autoren stellten fest, dass es zu einer guten Textdiagnose sowohl einer beschreibenden als auch einer bewertenden Perspektive bezüglich formaler und inhaltlicher Aspekte bedarf. Feser & Höttecke (2016) ergänzten durch ihre Untersuchungen im Physikunterricht die Qualitätsmerkmale auf der fachlichen Seite: der Rahmenbau der Erklärung (Wissenschaftlichkeit und Einbindung von Fachwissen und Belegen), die Rolle von Evidenzbezügen (Gültigkeit und Eignung der Belege), die Tiefe der Erklärung (Vollständigkeit und Komplexität des Arguments) und die Konsistenz der Erklärung (Relevanz und fachliche Richtigkeit von Belegen). Qualitätsmerkmale auf der sprachlichen Seite stellen hierbei die Lexik und Semantik und der damit einhergehende Fokus auf einen verwendeten Fachwortschatz und die Syntax und Stilistik mit der Anforderung nach sprachlich verdichteten und unterschiedlich verbundenen Sätzen sowie

stilistisch tendenziell distanzierter Sprache, dar (ebd.).

In einer Studie von Weiss et. al. (2013) zur Korrektur von Abiturklausuren durch Geographielehrkräfte werden Probleme der Trennschärfe zwischen Beurteilungskriterien inhaltlicher und sprachlicher Art thematisiert und damit darauf aufmerksam gemacht, dass es für eine valide und erfolgreiche Textkorrektur an klaren Unterteilungen und Beschreibungen von Kategorien, welche Kompetenzen und Leistungen abzeichnen, bedarf. Mit dem Fokus auf den Vorgang der Korrektur konnte hier festgestellt werden, dass im Schnitt am häufigsten Ausdruck-, Rechtschreibungs- und Zeichensetzungsfehler von den Lehrkräften erkannt wurden. Die Schreibrichtigkeit wird mit Korrekturzeichen zur Grammatik, dem Satzbau, allgemeinen Streichungen und Einschüben erfasst und macht durchschnittlich über alle korrigierte Klausuren 71% der markierten Fehler aus (Weiss et. al, 2013:13). Es konnten bei der Untersuchung der korrigierten Abiturklausuren Defizite bei der Bewertung und Korrektur von Argumentationsfehlern festgestellt werden, dies kann auf die nicht klar definierte Trennschärfe zwischen den Aspekten der Textorganisation und Argumentation zurückzuführen sein (Weiss et. al, 2013:13).

Der aktuelle Forschungsstand bezüglich der Schüler*innentextbeurteilung durch Lehrkräfte im Fachunterricht hat sich insbesondere darauf konzentriert, welche Kriterien die Lehrkräfte zur Beurteilung verwenden. Dabei wird deutlich, dass Fachlehrkräfte sowohl die Beurteilung fachspezifischer als auch sprachlicher Kriterien beherrschen müssen. Von den vorgestellten Studien konzentriert sich dennoch keine auf die Beurteilung argumentativer Texte. Aus der theoretischen Einführung und dem aktuellen Forschungsstand in den Fachdidaktiken lassen sich jedoch bereits einige allgemeine Beurteilungskriterien ableiten. Bezogen auf argumentative Texte wird dies im Weiteren genauer ausgeführt.

Kriterien zur Beurteilung argumentativer Schüler*innentexte

Werden Argumentationen vor einem strukturellen Hintergrund beurteilt, wird sich in den Fachdidaktiken meist an der Argumentationsstruktur nach Toulmin (1996) orientiert. Budke und Kuckuck (2017:25) merken hierzu an: „Die Qualität dieser Elemente von Argumentation kann allerdings nur vor dem jeweils unterschiedlichen fachlichen Hintergrund der Schulfächer beurteilt werden. Hinzu kommen fachspezifische Gütekriterien“. Im Geographieunterricht gestalten sich dabei gute Argumentationen als ergebnisoffen, multiperspektivisch, sowohl räumliche als auch zeitliche Aspekte einbeziehend und komplex begründet.

Weiss et. al (2013) beschreiben den Anspruch des Faches Geographie und den damit einhergehenden schriftlichen Leistungen als, „dass zum Präsentieren der fachwissenschaftlichen Inhalte eine qualitativ befriedigende Textstruktur, Argumentation und Verwendung von Fachausdrücken explizit erforderlich sind“ (Weiss et. al, 2013:13). Die Bewertungsvorgaben bei dem Zentralabitur für Geographie in Nordrhein-Westfalen bieten Lehrkräften beispielsweise Bewertungsorientierung. Werden diese genauer betrachtet, kann der Prüfling hierbei maximal 80 Punkte für die inhaltlichen Anforderungen bekommen und 20 für die Darstellungsleistung (Weiss et. al, 2013:12). Diese Leistung umfasst folgende fünf Kriterien: Textstrukturierung (5 Punkte), Argumentation, Beleg, Fachsprache (je 4 Punkte) sowie Schreibrichtigkeit (3 Punkte) (KMK, 2005). Bei der Argumentation als Darstellungsleistung wird

fokussiert, dass der Prüfling „beschreibende, deutende und wertende Aussagen schlüssig“ (Weiss et al., 2013:12) aufeinander bezieht. Beim „Beleg“ geht es um angemessene und korrekte Nachweise für Aussagen.

Um diese Kriterien zu untersuchen, bedient sich die Forschung unterschiedlicher wissenschaftlicher Ansätze zur Argumentationsanalyse. Häufig werden strukturelle Argumentationsanalysen durchgeführt (u.a. Lam et al., 2018; Abdollahzadeh et al., 2017; Stapleton & Wu, 2015; Riemeier et al., 2012; Simon et al., 2006; Basel et al., 2013; Zohar & Nemet, 2002; Chase, 2011; Knudson, 1992). Diese werden zur Analyse der Vollständigkeit und Komplexität von Einzelargumenten eingesetzt (z. B. Basel et al., 2013; von Aufschnaiter et al., 2008; Chase, 2011). Dabei wird sich in der Regel an den von Toulmin (1996) definierten Einzelementen eines Arguments orientiert. Teilweise steht der quantitative Aspekt mit der Anzahl der dargelegten Argumente im Vordergrund und wird als Qualitätsmerkmal gewertet (z.B. Benetos/ Betrancourt 2020; Basten et al. 2017). Mit holistischen Verfahren hingegen werden ganzheitliche Textbewertungen durch unterschiedliche Rater*innen durchgeführt, welche auch Bewertungskriterien berücksichtigen, wobei deren Gewichtung nicht immer definiert ist. Dabei geht es darum, Texte in Gruppen unterschiedlicher Qualität einzuteilen, (vgl. Studien von Knudson, 1992; Allen et. al, 2018; Chase, 2011, Abdollahzadeh et. al., 2017). Bei inhaltlichen Argumentationsanalysen wird der Fokus auf die fachliche Qualität gelegt. Die Qualität der Argumentationselemente lässt sich durch Gütekriterien wie die Relevanz der Schlussfolgerungen, Eignung der Geltungsbeziehungen und Gültigkeit der Belege (Kopperschmidt, 2000, Stapleton/ Wu, 2015), die Validität (Zohar/Nemet, 2002), die Passung der Argumente und fachlicher Richtigkeit beim Materialbezug (Sandoval, 2005) oder der Bezug auf wissenschaftliche Modelle (Böttcher/Meisert, 2011) bestimmen. Zuletzt nutzen linguistische Argumentationsanalysen Kriterien wie typische sprachliche Handlungen bei Argumentationen (Textprozeduren) (Andrews, 2016) oder die Berücksichtigung der Textsorte, Widerlegung von Gegenargumenten durch sprachliche Mittel und die Formulierung der eigenen Meinung (Quasthoff/ Domenech, 2016:32-33).

Dazu diskutieren Budke et. al. (2020a) die obigen Argumentationsanalysemethoden und -kriterien und entwickelten einen interdisziplinären theoriegeleiteten Kriterienkatalog für die wissenschaftliche Analyse von Schüler*innentexten¹ (113f). Dabei werden Kriterien mehrerer Ansätze genutzt, um eine ganzheitliche Beurteilung der Texte zu ermöglichen und die Stärken der einzelnen Verfahren zu nutzen und um Schwächen anderer Ansätze auszugleichen.

Im Nachfolgenden werden, die in der Literatur diskutierten Beurteilungskriterien für die vorliegende Studie zusammengefasst.

1 Link zu Analyseraster: <https://geodidaktik.uni-koeln.de/multimedia/analysebogen-zur-beurteilung-von-schriftlichen-argumentationen>

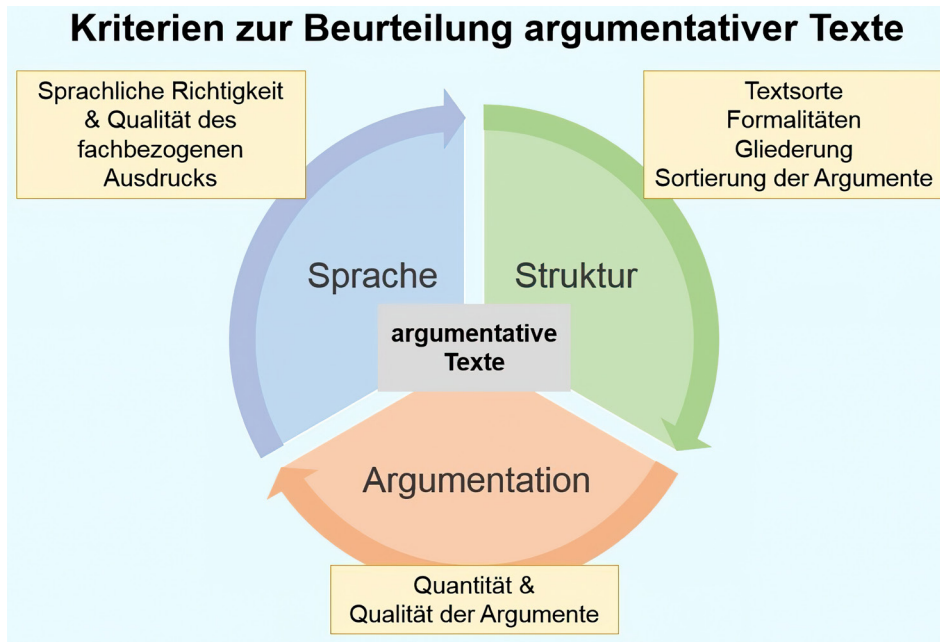


Abbildung 2 Kriterien für die Beurteilung argumentativer Texte (Quelle: eigene Darstellung)

Abbildung 2 verdeutlicht die grundlegenden Kriterien, die bei der Beurteilung eines argumentativen Schüler*innentextes auf der Grundlage der wissenschaftlichen Literatur durch Lehrkräfte berücksichtigt werden sollten. Die Oberkategorie ‚Sprache‘ umfasst unter anderem die Kriterien der korrekten Rechtschreibung, der angemessenen Grammatik und des Ausdrucks sowie die sachgemäße Verwendung von Fachtermini, welche ein gelungener argumentativer Text aufweisen sollte (Budke et al., 2020a). Des Weiteren ist eine Hauptkategorie die ‚Argumentation‘, in der die Quantität und Qualität der Argumente erfasst wird. Die dritte Oberkategorie ist die ‚Textstruktur‘ (Abb. 2). Dabei spielt insbesondere die Einhaltung der Textsorte (sowohl formell als auch sprachlich) sowie die grobe Gliederung in eine Einleitung, einen Hauptteil und einen Schluss, eine Rolle. Im Rahmen der Methodenbeschreibung wird das Modell präziser operationalisiert.

Methode

Tabelle 1 verdeutlicht die Forschungsfragen, die verwendeten Methoden der Erhebung und Auswertung, denen in der Studie nachgegangen wurde.

Forschungsfrage	Forschungsaspekte	Methodik
<i>Wie gehen Geographielehrkräfte bei der Beurteilung von argumentativen Schüler*innentexten vor und wie begründen sie ihr Vorgehen?</i>	Beurteilungsstrategien, Beurteilungsverhalten und Begründungen; Vergleich Interviewaussagen & Beurteilungsverhalten	Videoanalyse der Textbeurteilung anhand Analyseraster (vgl. Tabelle 2;3;4); Auswertung des Leitfadenterviews per qualitativer Inhaltsanalyse nach Mayring (2019) (MAXQDA)
<i>Welche Kriterien nutzen Geographielehrkräfte zur Beurteilung argumentativer Schüler*innentexte?</i>	Beurteilungsdauer; Anzahl der beurteilten Kriterien	Videoanalyse der Textbeurteilung anhand der Analyseraster (vgl. Tabelle 2;3;4); Zählung der berücksichtigten Kriterien
<i>Welche Methoden verwenden Geographielehrkräfte zur Beurteilung argumentativer Schüler*innentexte?</i>	Beurteilungs- & Korrekturmethode	Analyse der schriftlichen Beurteilungen der Lehrkräfte

Tabelle 1 Darstellung der Forschungsfragen & zugehörigen Erhebungs- & Auswertungsmethoden (Quelle: eigene Darstellung)

Die Erhebung erfolgte mit zehn Geographielehrkräften per Zoom-Call in Form von leitfadengestützten Interviews, welche per Bildschirmrecording aufgezeichnet, transkribiert und per qualitativer Inhaltsanalyse nach Mayring (2010) mit Hilfe der Kodierungssoftware MAXQDA sowie quantitativ ausgewertet wurden. Die Interviews umfassten Fragen zur Textbeurteilung, Feedback und Förderung: Die Lehrkräfte wurden zur Selbsteinschätzung ihrer Beurteilungsfähigkeiten argumentativer Schüler*innentexte und wo sie sich diese angeeignet haben, befragt. Des Weiteren beschrieben sie, welche Instrumente sie zur Schüler*innentextbeurteilung verwenden und welche Kriterien ein gelungener argumentativer Text aufweisen sollte. Hierbei ist zu erwähnen, dass während der Erhebung der Fachterminus ‚Korrektur‘ verwendet wurde.

Im Anschluss an die Interviewfragen bekamen die Lehrkräfte einen anonymen, authentischen Schülertext aus der 8. Klasse eines Gymnasiums, welchen sie korrigieren sollten. Der Schüler schrieb eine argumentative Stellungnahme zu einem lokalen Raumnutzungskonflikt. Dazu hatte der Schüler unterschiedliche kontinuierliche und diskontinuierliche Informationsmaterialien² zur Verfügung. Für den Text hat der Schüler 37 Minuten Zeit investiert und ihn handschriftlich formuliert. Für die Interviews mit den Geographielehrkräften und die Textbeurteilung wurde dieser abgetippt.

Die Lehrkräfte hatten nun die Aufgabe: *„Korrigieren Sie bitte den nachfolgenden argumentativen Schülertext. Bitte kommentieren Sie gleichzeitig verbal Ihre Gedanken und Ihre Vorgehensweise (Wie gehen Sie vor? Was korrigieren Sie? Warum?)“*. Als Hintergrundinformation wurde den Lehrkräften die Klassenstufe des Schülers genannt. Die Beurteilung fand innerhalb eines Word-Dokuments anhand der Kommentarfunktion und der Funktion ‚Änderung nachverfolgen‘, wodurch das Vorgehen nachvollziehbar wurde, statt. Dies wurde per Bildschirmrecording aufgenommen und durch das Laut-Denken (Konrad, 2020) der Lehrkräfte begleitet. Anhand dessen wurde der Verlauf der Beurteilung durch die Lehrkräfte festgehalten und durch ihre verbalen Kommentare untermauert. Bei der Datenerhebung ermöglichte die Videographie und Transkripte des Lauten Denkens detailliertere Einblicke in die Denkprozesse, welche durch alleinige Beobachtung entfallen würden (Heine & Schramm, 2016).

Bei der Auswertung werden die Interviewantworten vor der Textbeurteilung mit der Aufzeichnung des tatsächlichen Vorgangs verglichen und untersucht, ob die Lehrkräfte ihre genannten Kriterien während ihrer Beurteilung tatsächlich berücksichtigen. Der Vergleich soll dazu dienen, herauszufinden, wie bewusst, sicher und routiniert die Lehrkräfte bei der Schüler*innentextbeurteilung vorgehen.

2 Link zu Unterrichtsmaterialien: <https://geodidaktik.uni-koeln.de/multimedia/raumnutzungskonflikt-innerer-gruenguertel-koeln>

Darüber hinaus wurde untersucht, welche Methoden die Lehrkräfte verwenden, um die Texte zu beurteilen – textimmanente Korrekturen, Randkommentare oder mündliche Beschreibungen. Die Ergebnisse können Routinen, Hürden und Unsicherheiten aufdecken, über welche Beurteilungskriterien Schüler*innen am Ende Feedback bekommen.

Neben der qualitativen Auswertung der Interviewaussagen und der Textbeurteilung fand im Weiteren eine quantitative Analyse statt, um zu untersuchen, welche Kriterien die Geographielehrkräfte häufiger und weniger oft thematisiert haben. Als Auswertungsmethode wurden drei Beurteilungsraster argumentativer Schüler*innentexte entworfen, welche die drei bereits vorgestellten Oberkategorien (Abb. 2) aufgreifen und operationalisieren (Tab. 2-4). Die Beurteilungsraster umfassen sowohl induktive Kategorien, welche sich im Verlauf der Videosequenzen als relevant herausgestellt haben als auch deduktive Kategorien, welche aus der Theorie entwickelt wurden (Abb. 2).

Sprachliche Kriterien				
Kriterium	Kürzel ¹	Erklärung	Zitate aus Schüler*innentext	Handlungen der Lehrkräfte (LD = Laut Denken, A = Anmerkung Textrand, K = textimmanente Korrektur)
Rechtschreibung	R	Rechtschreibung (KMK, 2005; Weiss et. al., 2013). Umfasst auch gendern	<i>pickniecken</i>	K = picknicken (wBGesU01)
Zeichensetzung	Z	Kommasetzung, Interpunktion (Budke et. al, 2020; Weiss et. al., 2013)	<i>„(Zitat) sagt der [...]</i>	LD = „Wörtliche Rede muss noch beendet werden“ (wBGesU01) K = „(Zitat)“, sagt der [...] (mTGesU0)
Grammatik	Gr	Syntax, Konjunktionen, Textverweise (KMK, 2005, Budke et. al, 2020)	<i>Beispielsweise um dort ihre Hunde auslaufen zu lassen.</i>	A = Satzbau (wNGymU0)
Tempus	T	Teilbereich der Grammatik, korrekte Zeitform (Präteritum, Präsens, Futur etc.) (KMK,2005:8)	-	-
Modus	M	Der Modus zeigt die Haltung des/ der Verfasser*in, ob es sich um eine Tatsache, Möglichkeit, Notwendigkeit, Wunsch oder Befehl handelt (Indikativ, Konjunktiv, Imperativ) (Crystal, 2008; KMK,2005:8)	<i>Außerdem würden Spieler Granulat verteilen, was der Umwelt schadet.</i>	K = „[...] schaden könnte“ (mBGesU)
Ausdruck	A	Sammelkategorie semantischer & textsortenspezifischer Fehlerarten: Wörter, Wortfolgen, Sätze, die missverständlich, mehrdeutig, kontextunangemessen und/oder in ihrer Lexik, Wortverbindung, Stilistik unpassend sind (Henke, 2022; KMK,2005:8; Feser & Höttecke, 2016; Weiss et. al., 2013)	<i>Bürger*innen meinen das Steuergelder für dieses Projekt verschwendet werden.</i>	LD = „Meinen find ich immer so ein schwieriges Verb. Da würde ich jetzt als Verbesserungsvorschlag vermuten“ A = A: besser: vermuten (wBGesU01)
Textprozeduren	TP	Sprachliche Mittel wie Verknüpfungswörter, Textstrukturierungselemente, Konjunktionen, Adverbien, Konnektoren, Adverbialsätze, die dazu dienen Argumente miteinander zu verknüpfen, strukturieren und Textkohärenz zu begünstigen. Sie unterstützen sprachlich beim Positionieren & Perspektivieren, Konzedieren, Kontrastieren, Begründen & Schlussfolgern (Feilke, 2014; Schüler, 2017; Andrews, 2016; Budke et. al, 2020)	<i>Ich denke, dass es nicht richtig wäre, [...]</i>	A = „Verwende eher die Formulierung ‚Meiner Meinung nach‘“ (mBGesU)
Lexik/ Fachtermini	FW	Korrekte Verwendung von Fachbegriffen (KMK, 2005, Tajmel, 2011; Feser & Höttecke, 2016; Busch & Ralle, 2013; Weiss et. al. 2013)	<i>[...] weil der Boden sonst versiegelt wäre, wodurch Tiere sterben würden.</i>	LD = „Das ist ja schon ein geographischer Begriff, sodass er erklären sollte, was er damit meint“ (mNGesU)
Sonstiges	SP-	Beispiel:		-

Tabelle 2 Beurteilungskriterien der „Sprachlichen Komponenten“ (Quelle: eigene Zusammenstellung)

Anmerkung: Die Beurteilungskürzel zur Rechtschreibung, Zeichensetzung, Tempus, Modus und Ausdruck sind aus dem gängigen Korrekturzeichensystem von Kepplin (2003:144) übernommen sowie an den Abiturvorgaben für Nordrhein-Westfalen orientiert. Dieser fokussiert sich ausschließlich auf sprachliche Beurteilungskriterien und unterteilt das Kriterium der „Grammatik“ in mehrere Unterkategorien (bspw. Genus, Präpositionen etc.). Darauf wird hier verzichtet, da der Korrekturfokus auf argumentativen und strukturellen Fehlern liegt.

Tabelle 2 zeigt sprachliche Kriterien, die bei der Schüler*innentextbeurteilung relevant sind. Die Rechtschreibung, Zeichensetzung, Grammatik, Tempus, Modus und Ausdruck zählen dabei zu den allgemeinsprachlichen Kriterien, die interdisziplinär und fächerübergreifend gelten. In den Abiturvorgaben der KMK fallen diese Aspekte gesammelt unter den Punkt der ‚Sprachrichtigkeit‘, was so viel bedeutet wie: „Der Prüfling schreibt sprachlich richtig (Grammatik, Orthographie, Zeichensetzung) sowie stilistisch sicher“ (KMK, 2005, vgl. Weiss et. al., 2013:12). Besonderer Fokus des schriftlichen Argumentierens im Geographieunterricht liegt seitens der sprachlichen Ebene bei der korrekten Verwendung fachsprachlicher Termini und passender Textprozeduren im Argumentationszusammenhang.

Argumentation				
Kriterium	Kürzel	Erklärung	Zitate aus Schüler*innentext	Handlungen der Lehrkräfte (LD = Laut Denken, A = Anmerkung Textrand, K = textimmanente Korrektur)
Strittigkeit	AS	Darstellung des Konflikts (Toulmin, 1996)	<i>Die Einwohner, die in der Nähe des Grüngürtels wohnen wollen den Grüngürtel nutzen. Beispielsweise um dort ihre Hunde auslaufen zu lassen. Die Spieler, die nah am Grüngürtel wohnen wollen dann aber dort spielen.</i>	A= „Hier passen die Sätze nicht aufeinander. Du musst die Problematik genauer benennen.“ (wNResU2)
Positionen der Akteur*innen	AP	Nennung & Beschreibung (verschiedener) Positionen von Akteur*innen (Toulmin, 1996; Budke & Kuckuck, 2017)	<i>Bürger*innen meinen das Steuergelder für dieses Projekt verschwendet werden.</i>	A = „Werde hier konkreter: Gibt es eine konkrete Initiative, die Position ergreift?“ (mAGesU)
Schüler*innenpositionierung	AEP	Argumentationsstruktur: eigene Position (Quasthoff/ Domenech, 2016); Schlussfolgerung (Toulmin, 1996)	<i>Zusammenfassend ist zu sagen, dass Sie den 1. FC Köln nicht erlauben sollen, dass Trainingsgelände zu erweitern.</i>	A= „Wo ist deine Meinung?“ (wAGesU02)
Komplexität	AK	Multiperspektivität (Feser & Höttecke, 2016; Budke & Kuckuck, 2017); angemessene Quantität an Pro- & Kontraargumenten (Benetos/ Betrancourt 2020; Basten et al. 2017)	<i>Dieses Kriterium wird auf den gesamten Text bezogen.</i>	LD = „ich hätte [...] mir mehr Argumente gewünscht, die auch dafürsprechen. Das ist relativ einseitig.“ (mNIGYU0)
Begründung (Beleg & Geltungsbeziehung)	AB	Argumentationsstruktur: Wissenschaftlichkeit (Feser & Höttecke, 2016; Toulmin, 1996)	<i>Schon seit 1926 trainieren die Profis vom 1. FC Köln im Grüngürtel sagt der Vorstand des 1. FC Kölns. Ich Persönlich finde dieses Argument nicht überzeugend.</i>	A = „Warum nicht überzeugend?“ (wAGesU01)
Bedingungen (Raum/ Zeit)	IR/ IZ	Bedingungen, unter denen das Argument gilt (DGfG, 2020; Budke & Kuckuck, 2017)	<i>Die Bürger*innen und der 1. FC Köln streiten sich darüber, ob sie auf dem Grüngürtel eine Erweiterung der Sportplätze bauen.</i>	LD = „Ich habe keine Vorstellung von der Fläche“ A = „Angabe zu ungenau, um den Grüngürtel welcher Stadt handelt es sich?“ (wAGesU01)
Eignung	AG	Argumentationsqualität: Geltungsbeziehung, Logik zwischen Meinung und Beleg (Feser & Höttecke, 2016; Stapleton/ Wu, 2015)	<i>Der 1. FC Köln will das Sportgelände am Grüngürtel bauen, weil sich dort alle Sportplätze an einem Ort befinden/ befänden und es deshalb ein geeigneter Platz wäre</i>	LD = „die Begründung des Schülers passt nicht“ A = „eigene Argumente nicht überzeugend. Führe weitere mögliche Argumente auf“ (mTGeSU0)
Relevanz	AR	Argumentationsqualität: These/ Schlussfolgerung, Passung der Schlussfolgerung zur Problemfrage (Feser & Höttecke, 2016; Stapleton/ Wu, 2015)	<i>Die Einwohner, die in der Nähe des Grüngürtels wohnen wollen den Grüngürtel nutzen. Beispielsweise um dort ihre Hunde auslaufen zu lassen. Die Spieler, die nah am Grüngürtel wohnen wollen dann aber dort spielen.</i>	LD = „Hier finde ich die Argumentation so ein bisschen redundant, weil das ist klar.“ A= „Achtung Paraphrasenalarm! Diese Aussagen sind von vorne rein klar und in der Argumentation eher schwächer anzusetzen.“ (mAGesU)
Gültigkeit/ Validität	AGÜ	Argumentationsqualität: fachliche Richtigkeit des Belegs (Feser & Höttecke, 2016; Zohar/Nemet, 2002) & Ausführlichkeit	<i>Außerdem würden Spieler Granulat verteilen, was der Umwelt schadet.</i>	LD = „Ja Granulat kann ich nicht einschätzen. Ob da Granulat verteilt wird. Ich stelle mir jetzt einfach so einen Sportplatz vor, aber ich weiss nicht, was die da bauen wollen“ (wNResU2)
Materialbezug	MB	Impliziter/ expliziter Materialbezug, Synthese mehrerer Ausgangsquellen (Schüler, 2017), materialgestütztes Schreiben (Feilke et. al., 2016; Sandoval, 2005)	<i>Der 1. FC Köln will jedoch alles selbst finanzieren und eine Bushaltestelle einrichten, damit der Verkehr nicht steigt.</i>	LD = „Also die Frage ist natürlich, was die vorher an Materialien bekommen haben [...] Dann würde ich Leitfragen zum Material wie 'Arten Kosten aus?' ...'Lärmbelästigung' oder sowas stellen“ (mTGeSU0)
Sonstiges	A-	Beispiele: Kompromissfindung; Gewichtung der Argumente	<i>Zusammenfassend ist zu sagen, dass Sie den 1. FC Köln nicht erlauben sollen, dass Trainingsgelände zu erweitern.</i>	A= „Alternativen/ Kompromisse aufzeigen“ (mTGeSU0) A = „Hier nochmal schreiben, was dein überzeugendstes, stärkstes Argument ist“ (wAGesU01)

Tabelle 3 Beurteilungskriterien im Bereich „Argumentation“ (Quelle: eigene Zusammenstellung)

Tabelle 3 stellt Teilkriterien für eine gelungene Argumentation dar. Darunter fallen unter anderem, die Strukturkriterien eines Einzelarguments (für die Praxis umbenannt und zusammengefasst in: Strittigkeit, Beleg, Positionen/ eigene Position), angelehnt an Toulmin (1996) sowie gängige Qualitätskriterien (Relevanz, Eignung, Gültigkeit, Komplexität) (Kopperschmidt, 2000). Zeit- und Raumbedingungen, Akteur*innen und Materialbezug (Budke et al., 2020a) wurden in der Kategorie ‚Argumentation‘ zugeordnet, da hier alle inhaltliche Aspekte des Textes mit bewertet werden.

Textstruktur/ Textsorte				
Kriterium	Kürzel	Erklärung	Auszüge aus Schüler*innentext	Handlungen der Lehrkräfte (LD = Laut Denken, A = Anmerkung Texttrand, K = textimmanente Korrektur)
Textsorte (z.B. Briefform)	ST	Strukturelemente wie Absätze, angemessene Stilistik (KMK, 2005); Textsorte (Quasthoff/ Domenech, 2016)	<i>Das wollen die Bürger*innen nicht, weil sie nicht mehr picknicken, spazieren, joggen, ... können.</i>	LD = „zu umgangssprachlich für einen Leserbrief“ (wBGesU01) K = picknicken, spazieren, joggen usw. können. (mNiGymU0) A = „Bitte keine ... in Leserbrief verwenden. Besser etc.“ (wBGesU01)
Einleitung	SE	Konfliktbeschreibung, Einführung, Textstruktur (Weiss et. al., 2013; KMK, 2005)	<i>Die Bürger*innen und der 1. FC Köln streiten sich darüber, ob sie auf dem Grüngürtel eine Erweiterung der Sportplätze bauen. Der 1. FC Köln möchte das natürlich, damit ihre Spieler besser trainieren können.</i>	LD = „Eigentlich hätte ich mir das allgemeiner gewünscht am Anfang, dass es eben verschiedene Positionen gibt“ A = „Hier bitte noch etwas allgemeiner einleiten“ (mNiGymU0)
Hauptteil	SH	Positionen, Komplexität, Textstruktur (Weiss et. al., 2013; KMK, 2005)	<i>Durch Flucht von der Sportanlage würden Fledermäuse und viele andere Insekten gestört werden. Bürger*innen meinen das Steuergelder für dieses Projekt verschwendet werden.</i>	LD = „...das ist der Fall von ganz vielen Argumenten aneinandergereiht, aber die hängen nicht thematisch miteinander zusammen“ A = „nicht verschiedene Argumente aneinanderreihen, sondern jedes Argument für sich ausführen“ (wBGesU01)
Schluss	SCH	Gesamturteil & eigene Meinung oder Kompromiss, Textstruktur (Weiss et. al., 2013; KMK, 2005)	<i>Zusammenfassend ist zu sagen, dass Sie den 1. FC Köln nicht erlauben sollen, dass Trainingsgelände zu erweitern.</i>	LD = „Naja das Urteil ist jetzt natürlich deutlich zu kurzgefasst und widerspricht teilweise der vorangegangenen Argumentation“ A = „Urteil zu kurzgefasst“ (mBGesU)
Sonstiges	S-	Beispiel:		

Tabelle 4 Beurteilungskriterien im Bereich der „Textstruktur/ Textsorte“ (Quelle: eigene Zusammenstellung)

Tabelle 4 zeigt Textgliederungsaspekte auf. Eine gelungene Struktur innerhalb argumentativer Texte kann zur besseren Verständlichkeit, Kohärenz, Überzeugungskraft und effektiven Kommunikation (vgl. Graff & Birkenstein, 2010; Toulmin, 1996) beitragen. Hierzu ist von Belang zu untersuchen, ob die Lehrkräfte, die in der Aufgabenstellung geforderte Textsorte (Brief) und die dazugehörigen Kriterien (vereinfacht: Anrede, Anliegen, Grußwort inklusive Zeichen- und Absatzregel) erkennen und beurteilen.

Proband*innenauswahl

Die Proband*innenauswahl wurde im Sinne des *theoretical samplings* (Glaser & Strauss, 1967) durchgeführt. Da es sich um eine explorative, qualitative Studie handelt, wurde angestrebt, möglichst kontrastäre Geographielehrkräfte zu finden, um vielfältige Perspektiven aufzugreifen (Dimbath, Ernst-Heidenreich & Roche, 2018). Daher variierten das Geschlecht, das Alter, die Berufserfahrung, die Schulform an der unterrichtet wird und die Zweifächer der Proband*innen. Für die Interpretation und den Rückbezug zu den Aussagen der Proband*innen, wurden diese mit Codes ausgestattet. Diese beinhalten Informationen über das Geschlecht der Lehrkraft (m/w/d), das jeweilige Bundesland der Schule (Ni= Niedersachsen, N= Nordrhein-Westfalen, B= Brandenburg,

T = deutsche Schule in Tokyo), die Schulform (Gym = Gymnasium, Ges = Gesamtschule, Rea = Realschule), ob sie Sek I und/ oder Sek II unterrichten (U = Sek I, O = Sek II, UO = beides) sowie eine Nummerierung (1,2) bei denselben Eigenschaften.

Beschreibung der Ergebnisse

Interviewergebnisse: Einstellungen der Geographielehrkräfte zur Textkorrektur und -beurteilung

Die Lehrkräfte wurden in den Interviews zunächst gefragt, warum sie Schüler*innentexte korrigieren. Die meisten Proband*innen antworteten, dass sie die Schüler*innen dadurch in erster Linie individuell fördern wollen – vereinzelt wurde hierzu ergänzt, dass somit eine Rückmeldung über ihren aktuellen Lernstand miteinhergeht. Zwei Lehrkräfte verbanden damit zudem eine Möglichkeit einer positiven Lernerfahrung für die Schüler*innen: *„...sodass ich einfach finde, dass man als Grundlage für das Lernen das auch nutzen kann oder sich auch mal auf die Schulter zu klopfen und zu sagen ‚Hier in der Q1³ war’s noch so, aber jetzt kann ich es schon besser‘“* (wBGesUO1) – hierbei spricht die Lehrkraft aus der vermuteten Sicht des Schülers bzw. der Schülerin. Darüber hinaus erwähnten zwei Lehrkräfte, dass sie die Texte korrigieren, um die Schüler*innen zum Arbeiten zu motivieren und Anforderungen durchzusetzen: *„Ja das hört sich witzig an, aber das ist tatsächlich so eine kleine Maßnahme, damit die überhaupt mal ein bisschen ordentlich arbeiten“* (wNReaU2). Dieselbe Probandin ergänzt als Einzige jedoch auch, dass sie Texte einsammelt und korrigiert, um ihren eigenen Unterricht zu evaluieren.

Drei Lehrkräfte hinterfragten den Sinn und Mehrwert der Korrektur für die Schüler*innen. Hierbei wurde nicht grundsätzlich die Textbeurteilung hinterfragt, sondern die Art und Weise der Korrektur und der didaktischen Einbettung im Unterricht: *„Wie viel wert ist das eigentlich, wenn ich das Ding wiederkrieg und da steht überall nur rot, rot, rot, rot, rot, was alles falsch ist. Ist die Frage, ob das so motivierend ist und ob das alles immer so sinnvoll ist oder ob man sich nicht viel besser auf zwei oder drei Dinge konzentriert“* (mNGymUO2). Der Proband kritisiert hierbei die überwiegend textimmanente Korrektur, die sich ausschließlich auf die Erfassung von Fehlern fokussiert. In diesem Zuge ist es wertvoll zu erwähnen, dass die meisten Lehrkräfte in den Interviews erläuterten, dass sie sowohl Fehler als auch positive Aspekte des Textes in ihren Randnotizen erwähnen und damit den Begriff der ‚Korrektur‘ eher synonym zur ‚Textbeurteilung‘ sehen als das ausschließliche Korrigieren von Fehlern: *„Und..ja, das versuche ich auch so ein bisschen abzufedern, indem ich tatsächlich auch wertschätzende Kommentare schreibe“* (mTGesUO). Auffällig ist jedoch, dass in der tatsächlichen Beurteilung des Textbeispiels in der Studie lediglich eine Lehrkraft einen positiven Randkommentar formulierte.

Im weiteren Verlauf wurden die Lehrkräfte dazu befragt, in welchem Kontext sie sich ihre Korrekturfähigkeiten angeeignet haben. Fünf Lehrkräfte betonten, dass sie die Fähigkeiten in ihrem Zweitfach (meistens Deutsch) erworben haben. Drei Lehrkräfte betonten explizit, dass

3 Q1 bezeichnet die Qualifikationsphase 1, welche in NRW Gymnasien und Gesamtschule die Klasse 11 meint.

Korrekturkompetenzen im Studium nicht vermittelt wurden und lediglich eine Lehrkraft konnte widersprechen. Drei Lehrkräfte erklärten, dass sie Kenntnisse in dem Referendariat erworben haben. Alle Lehrkräfte berichteten, dass sie ihre Korrekturfähigkeiten in der Schule als Lehrkraft durch ‚learning-by-doing‘ vertieft hätten. Eine Lehrkraft resümiert: *„Ja das ist ein lebenslanger Prozess. Der beginnt sicherlich in der eigenen Schulzeit, dann im Studium und [...] dann habe ich früher Nachhilfe gegeben und joah natürlich jetzt 13 Jahre oder bzw. 15 Jahre waren es ja schon mit Referendariat, wo man das so lernt. Und vereinzelt sicherlich auch durch das eigene Schreiben von Arbeiten im Studium, da muss man sich ja auch selber korrigieren“* (mNiGymUO).

Fast alle Lehrkräfte beschrieben ihre Korrekturfähigkeiten als durchschnittlich bis sehr gut. Lediglich zwei Lehrkräfte konnten hierzu keine Einschätzung abgeben. Es wurde jedoch mehrfach von den Proband*innen betont, dass Unsicherheiten bei der Beurteilung argumentativer Texte bestehen und dies durch die Komplexität der Textsorte stets eine Herausforderung darstellt.

Darüber hinaus betonten insbesondere die Lehrkräfte mit dem Zweitfach Deutsch, dass sie dadurch bereits eine grundlegende Sicherheit bei der Korrektur von sprachlichen Aspekten, wie beispielsweise Rechtschreibung und Grammatik haben. Bis auf einer Lehrperson fällt die Anwendung dieses Korrekturkriteriums den Proband*innen leicht. Nachdem die Lehrkräfte ihre Selbsteinschätzung vorgenommen haben, beurteilten sie einen Schülertext. Wie sie vorgegangen sind und welche Kriterien zur Beurteilung argumentativer Schüler*innentexte genutzt haben, wird im Nachfolgenden dargestellt.

Ergebnisse aus der Textbeurteilung der Geographielehrkräfte

1. Wie gehen Geographielehrkräfte bei der Beurteilung von argumentativen Schüler*innentexten vor?

Zur Beurteilung des Schülertextes haben die Lehrkräfte zwischen 02:40 und 36:31 Minuten und im Durchschnitt 14:30 Minuten investiert. Dies zeigt eine große Spannweite von detaillierter Beurteilung mit textimmanenten Korrekturen, Streichungen, Ergänzungen und Randkommentaren bis hin zu holistischen Verfahren, bei denen die Proband*innen den Text lasen, ihren ersten Eindruck schilderten und ein Globalurteil fällten.

Vorbereitung auf Textbeurteilung	1. Textsichtung	2. Textsichtung
Erwartungshorizont/ Beurteilungskriterien erstellen Beschreibung: = 8	Sprache Beschreibung: = 7 Handlung: = 7	Sprache Beschreibung: = 1 Handlung: = 2
Aufgabenstellung lesen Beschreibung: = 7 Handlung: = 8	Struktur Beschreibung: = 5 Handlung: = 6	Struktur Beschreibung: = 1 Handlung: = 5
Materialien zur Aufgabe sichten Beschreibung: = 2 Handlung: = 4	Argumentation Beschreibung: = 2 Handlung: = 2	Argumentation Beschreibung: = 8 Handlung: = 5
	Textsichtung ohne Beurteilung Beschreibung: = 2 Handlung: = 3	Textsichtung ohne Beurteilung Beschreibung: - Handlung: -
	Gesamt Beschreibung: 10 Handlung: 10	Gesamt Beschreibung: = 8 Handlung: = 6

Tabelle 5 Darstellung des Beurteilungsverlaufs der zehn Geographielehrkräfte (Quelle: eigene Darstellung) Kommentar: „Beschreibung“ = Antwort zur Frage „Wie würden Sie bei der Textkorrektur vorgehen?“; „Handlung“ = tatsächliches Vorgehen. Nicht alle Geographielehrkräfte haben zwei Textsichtungen bzw. -beurteilungen vorgenommen und nicht alle haben jedes Kriterium korrigiert, deswegen variieren die „Gesamt“-zahlen.

Tabelle 5 zeigt die Herangehensweisen der Lehrkräfte bei der Schülertextkorrektur. Die Zahlen innerhalb der Tabelle stehen dabei für die Anzahl der Lehrkräfte. Bei der Auswertung wurde deutlich, dass sich die praktische Textbeurteilung in drei Hauptphasen unterteilen lässt: die Vorbereitung auf die Beurteilung, die erste Textsichtung und die zweite Textsichtung. Es wird innerhalb der Tabelle unterschieden zwischen der Beantwortung der Interviewfrage, wie die Lehrkräfte bei der Textbeurteilung vorgehen würden (Beschreibung) und der tatsächlichen Beurteilung, welche die Lehrkräfte vorgenommen haben (Handlung). Hierbei ist zu beachten, dass die Beschreibung der Lehrkräfte nicht immer mit ihrer Handlung übereinstimmt und dass teilweise Kriterien beschrieben wurden, die dann nicht beurteilt oder korrigiert wurden.

Bei acht Proband*innen liegen Differenzen zwischen beschriebener Handlung und tatsächlicher Durchführung vor – nur Wenigen war dies auch bewusst. So beschreibt eine Probandin: „*Ach jetzt fange ich gerade schon inhaltlich an*“ (wBGesUO1) entgegen ihres Planes den Text zunächst nur sprachlich zu korrigieren.

Aus der Darstellung geht hervor, dass die meisten Lehrkräfte in ihrem Unterricht mit Erwartungshorizonten arbeiten und im Vorhinein Beurteilungskriterien festlegen. Der Umfang des Interviews sah nicht vor, dass sie einen Erwartungshorizont erstellen sollten, deswegen beschränkt sich hier die Auswertung auf die „Beschreibung“. In den Interviews wurden Erwartungshorizonte von vier Lehrkräften explizit beschrieben. Sie nutzen diese größtenteils nur zur Beurteilung von Klausuren. Im Unterricht werden bei ihnen Texte meist frei beurteilt und mündlich besprochen. Drei Lehrkräfte betonten, dass sie im Vorfeld einer Schreibaufgabe, partizipativ mit den Schüler*innen Kriterien für den Text festlegen und somit eine Orientierung bieten oder zumindest ihre Kriterien im Vorfeld transparent kommunizieren.

Zusätzlich zu den Materialien würden sich die meisten Lehrkräfte die gestellte Aufgabenstellung anschauen. Die meisten Lehrkräfte gaben an, den Schülertext zweimal zu lesen bzw. zu beurteilen. Lediglich zwei Geographielehrkräfte handelten allerdings tatsächlich so. Die meisten Lehrkräfte sichteten den Text zweimal – lediglich drei nahmen in beiden Sichtungen schriftliche Textbeurteilungen oder -korrekturen vor. Insgesamt zeigt sich jedoch, dass das Vorgehen der Geographielehrkräfte stark variiert.

In beiden Textsichtungsdurchläufen haben alle Lehrkräfte strukturelle, neun Lehrkräfte sprachliche und lediglich sieben Lehrkräfte argumentative Kriterien im Text korrigiert. Davon wurden in der ersten Textsichtung überwiegend sprachliche und strukturelle Kriterien und in der zweiten Textsichtung argumentative Aspekte beurteilt. Die größte Diskrepanz zwischen der vorherigen Beschreibung und der tatsächlichen Handlung gibt es bei der Angabe der strukturellen Korrektur in der zweiten Textsichtung. Mit dem Blick auf die Spalte der Argumentation kann festgestellt werden, dass diese Proband*innen größtenteils angaben, dass sie argumentative Aspekte beurteilt haben, jedoch eigentlich die Struktur des Textes beurteilten oder/und korrigierten.

2. Welche Kriterien nutzen Geographielehrkräfte zur Beurteilung argumentativer Schüler*innentexte?

Zur Auswertung der Lehrkräfthandlungen, wurden die beurteilten Kriterien durch die Beurteilungsraster (vgl. Tabelle 2-4) zugeordnet und quantitativ ausgewertet. Die Abbildung 3 stellt die beurteilten Kriterien in den Oberkategorien Sprache, Struktur und Argumentation dar.

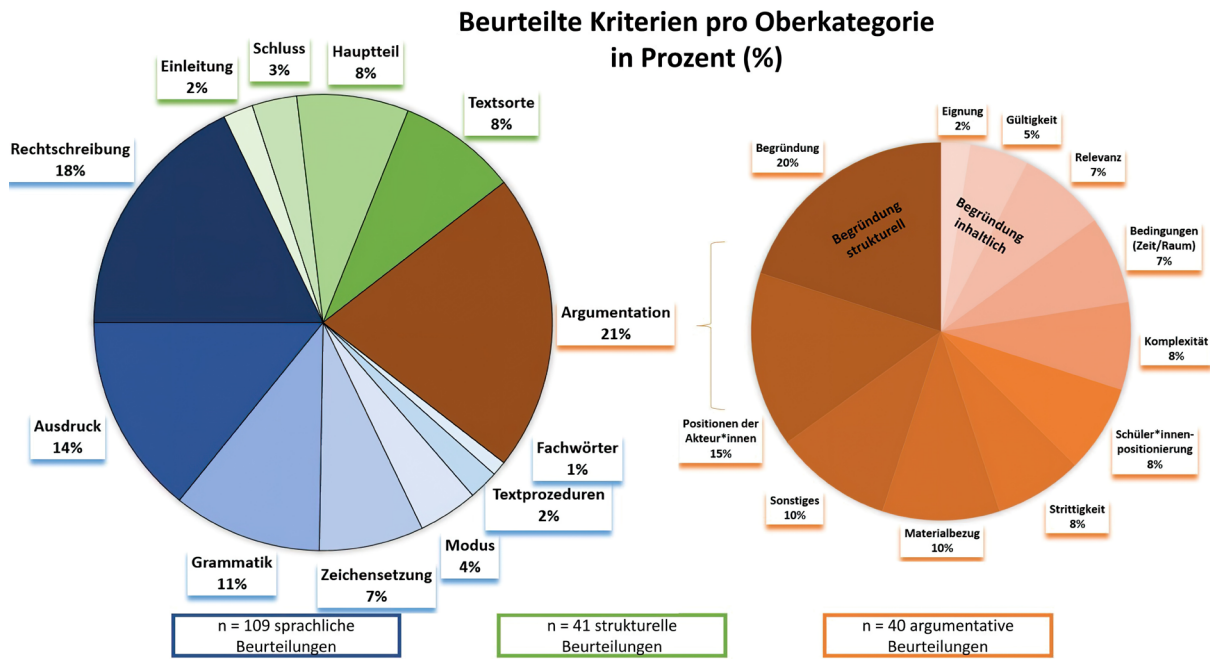


Abbildung 3 Beurteilungen zu den drei Oberkategorien (Sprache, Argumentation und Struktur, N =190 Beurteilungen) in Prozent (Quelle: Eigene Darstellung) Anmerkung: Die Kategorien ‚Tempus‘ und ‚Sonstiges‘ wurden nicht aufgeführt, da im Text keine Tempusfehler vorhanden waren und die Lehrkräfte keine Kriterien nannten, die unter die Kategorie ‚Sonstiges‘ gewertet wurden.

Insgesamt haben die 10 Proband*innen innerhalb des Schülertextes 190 Beurteilungen vorgenommen. Aus dem Kreisdiagramm (vgl. Abb. 3) wird deutlich, dass sich 57% (109 Beurteilungen) und damit über die Hälfte aller Korrekturen auf sprachliche Fehler beziehen. Sehr ausgeglichen ist das Verhältnis zwischen der Fehlerkorrektur struktureller Aspekte (41 Beurteilungen) und die der Argumentation (40 Beurteilungen).

Wenn sprachliche Aspekte beurteilt und/oder korrigiert wurden, dann größtenteils die Rechtschreibung (31%) und der Ausdruck (25%). Besonders gering fielen die Beurteilungen zum Fachwortschatz und zu den argumentativen Textprozeduren aus. In den leitfadengestützten Interviews hat die Mehrheit der Lehrkräfte sprachliche Kriterien für einen gelungenen argumentativen Text genannt: Ausdruck und Textprozeduren (je dreimal), sprachliche Variation und Rechtschreibung (je einmal). Kriterien wie die Verwendung von Fachwörtern oder eine angemessene Grammatik wurden nicht erwähnt.

Deutlich weniger Beurteilungen (40) wurden innerhalb der Oberkategorie ‚Argumentation‘ getätigt (Abb. 3). Im Gegensatz zu sprachlichen Beurteilungen zeigt sich hierbei eine gemischte Verteilung, wobei am häufigsten Defizite bei der Begründung beurteilt und/oder korrigiert wurden. Dennoch handelt es sich hierbei lediglich um 20% der argumentationsbezogenen Korrekturen (8 Beurteilungen).

Auffällig ist, dass die Lehrkräfte eher argumentationsstrukturelle Aspekte wie das Vorhandensein von Positionen, Materialbezug, Begründungen und der Strittigkeit beurteilen und Gütekriterien, welche sich eher mit der Qualität des Arguments beschäftigen – die Logik (Eignung: 1 Korrektur), fachliche Richtigkeit (Gültigkeit: 2 Beurteilungen) sowie die Relevanz (3 Beurteilungen) des Arguments – kaum beurteilen.

Randkommentare der Lehrkräfte kritisieren oftmals die fehlende Präzision und Begründung der Argumente. So schrieb eine Lehrkraft als Randkommentar: *„Bitte nicht verschiedene Argumente aneinanderreihen, sondern jedes Argument für sich ausführen (Argument nennen, erläutern, begründen, ggf. verdeutlichendes Bsp. nennen“* (wBGesUO1). Randkommentare zu qualitativen Gütekriterien der Argumentation (Eignung, Relevanz, Gültigkeit) wurden selten getätigt. Ein Beispiel umfasste die Kritik an der Eignung des Schlussfazits des Schülers: *„Urteil zu kurzgefasst und widerspricht sich z.T. mit der vorangegangenen Argumentation. Hier wäre ein Rückbezug auf die von dir geschilderten Umweltschäden nötig“* (Randkommentar mBGesU). Ebenso kritisierte ein Proband *„Relevanz unklar“* (Randkommentar mNiGymUO).

Zur Oberkategorie ‚Argumentation‘ hat jede Lehrkraft im Schnitt 3,3 Kriterien im Interview genannt. Darunter wurden am häufigsten die Schüler*innenpositionierung, die Begründung und die Komplexität als Kriterien für einen gelungenen argumentativen Schüler*innentext genannt. Im Rahmen der Komplexität wurde mehrfach genannt, dass die Anzahl der Argumente ausreichen sollte. Anschließend mit jeweils drei Nennungen pro Kategorie folgten die Darstellung der Strittigkeit, Positionen der Akteur*innen und Materialbezug. Vier weitere Nennungen beschrieben zwar keine konkreten Kriterien, jedoch dass ein argumentativer Text *„überzeugend“* und *„nachvollziehbar“* sein sollte. Insgesamt ist auffällig, dass den Lehrkräften die konkrete Benennung von Argumentationskriterien schwerfiel und oftmals wie in folgendem Beispiel nur grob erklärt werden konnte, worauf bei Schüler*innentexten geachtet wird: *„Das es nachvollziehbar ist, dass es nicht komplett an den Haaren herbeigezogen ist“* (wNReaU1).

Nahezu alle Lehrkräfte haben Textstrukturelemente als Kriterium für einen gelungenen argumentativen Text in den Interviews genannt – insgesamt 20-mal wurden Teilkriterien genannt. Die Einleitung (5 Nennungen), der Hauptteil und der Schluss (je 3 Nennungen) spielten für die Proband*innen die größte Rolle. Weitere Kriteriennennungen waren ungenauer und bezeichneten, dass ein guter argumentativer Text einen gelungenen Aufbau, Abschnitte und eine sinnvolle Gewichtung der Argumente braucht. Zwei Lehrkräfte nannten zusätzlich die Einhaltung der Textsorte. Die Auswertung der Interviewpassagen zeigt, dass die Proband*innen die strukturellen Kriterien im Gegensatz zu den Argumentationskriterien deutlich besser und sicherer benennen konnten und ihnen auch mehr Gewicht verliehen haben. Eine Probandin beschreibt dies expliziter: *„Also ich bin ein absoluter Fan von Struktur, weil ich glaube einfach, wenn man sich vorher gut überlegt, was man sagen will und in welcher Reihenfolge und dann gibt es ja auch so bestimmte Kniffe, was man wann sagt, damit es vielleicht besonders überzeugend ist“* (wNGymUO).

Anders als in den Interviews angegeben, haben die Lehrkräfte am häufigsten Anmerkungen und Korrekturen, welche die Textsorte des Briefes betreffen, vorgenommen (16 Beurteilungen).

Dies waren Kriterien wie fehlende Absätze: „Also ich muss mir den Text ja gar nicht durchlesen, also insgesamt...es gibt nirgendwo eine Leerzeile“ (mNGymUO2), defizitäre Grußform und Verabschiedung: „Adressatenbezug zu schwammig: Rückantwortwunsch fehlt“ (Randkommentar mNiGymUO) sowie ein für Briefe unangemessener Ausdruck. Im Hauptteil kritisierten die Lehrkräfte oftmals die Anordnung der Argumente. In der Einleitung hätten sich einige Lehrkräfte eine allgemeinere Einführung und im Schluss eine ausführlichere Beschreibung der eigenen Meinung des Schülers gewünscht.

Die Kriterien, welche die Lehrkräfte beurteilten, unterscheiden sich teils stark und liegen im Ermessen der Lehrkraft.

3. Welche Methoden verwenden Geographielehrkräfte zur Beurteilung argumentativer Schüler*innentexte?

Um zu überprüfen, inwiefern die Lehrkräfte die Kriterien (vgl. Tabelle 2-4) beurteilen, korrigieren oder lediglich erläutern, wurden die Handlungen durch die Lehrkräfte im Weiteren untersucht.

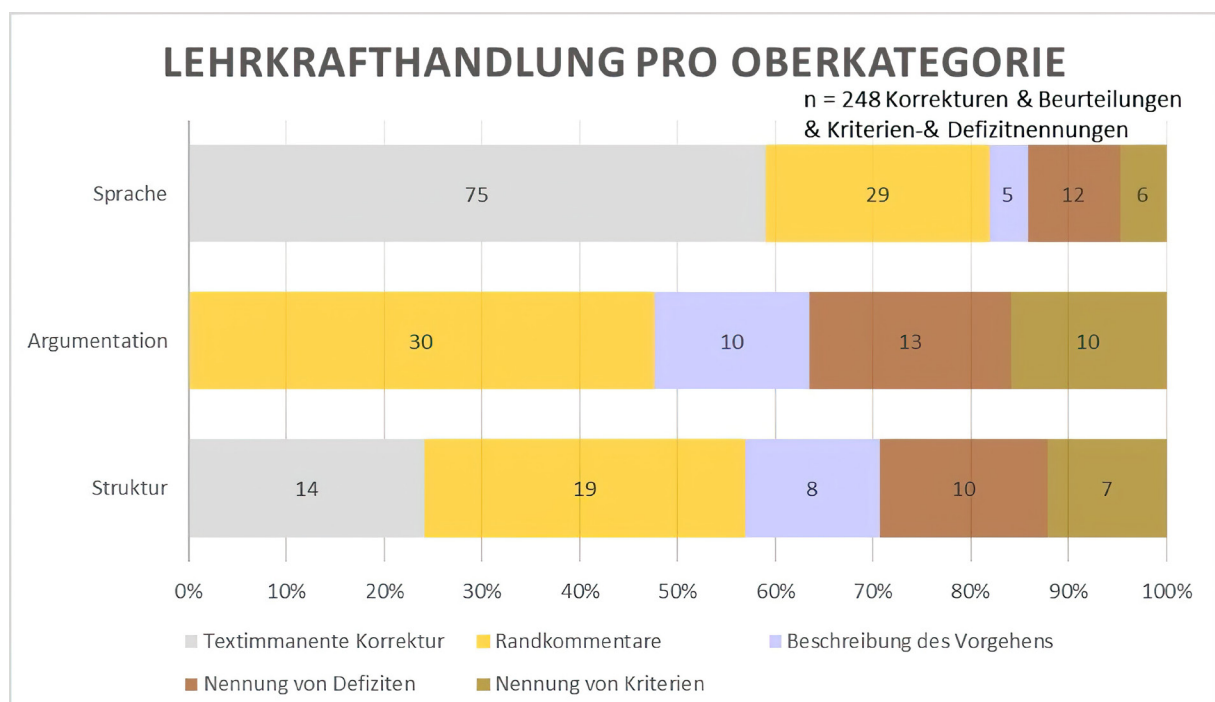


Abbildung 4 Darstellung der Lehrkrafthandlungen pro Oberkategorie (Quelle: eigene Darstellung) Textimmanente Korrektur (direkte Streichungen, Wortverbesserungen), Randkommentare und die Beschreibung des Vorgehens wurden als „Korrektur“ gewertet. Nennung von Kriterien & Defiziten wurde nicht als Korrektur gewertet und beschreibt Kriterien für einen gelungenen argumentativen Text, welche die Lehrkräfte parallel zu ihrer Korrektur mündlich geäußert, jedoch nicht korrigiert haben

Abbildung 4 zeigt, anhand welcher Methodik die Proband*innen ihre Beurteilung im Schülertext per textimmanenter Korrektur, Randkommentaren oder Beschreibung des eigenen Vorgehens vorgenommen haben. Teils haben die Lehrkräfte lediglich Defizite des Textes sowie Kriterien für einen gelungenen argumentativen Text geäußert ohne schriftliche Korrekturen vorzunehmen (Abb. 4). Wenn Randkommentare verwendet wurden, dann haben die Lehrkräfte oftmals Korrekturkürzel verwendet. Dies betrifft insbesondere die Korrektur von sprachlichen Aspekten (Bsp. R = Rechtschreibung).

190 schriftliche Korrekturen wurden insgesamt vorgenommen und 58 Aspekte lediglich mündlich kritisiert. Durchschnittlich hat jede Geographielehrkraft 19 Kriterien korrigiert und/oder beurteilt. Dabei gab es jedoch eine große Spannweite – vier Lehrkräfte haben maximal elf Fehler korrigiert und/oder beurteilt, fünf weitere 17 bis 33 und ein Proband 43. Auch bei den Kriterien- und Defizitnennungen gibt es starke Unterschiede.

Abbildung 4 zeigt, dass am häufigsten bei sprachlichen Fehlern textimmanent korrigiert wurde. Fast die Hälfte der Beurteilungen zu argumentativen Fehlern wurde dagegen durch Randkommentare festgehalten. Darüber hinaus nannten die Proband*innen insbesondere im Bereich der Argumentation 23-mal Defizite und Kriterien eines gelungenen argumentativen Textes, nahmen jedoch keine schriftliche Korrektur vor. Auf die Lehrkräfte bezogen, zeigten sechs Proband*innen textimmanente Korrekturen, acht Lehrkräfte Randkommentare und vier Personen korrigieren kaum und beschreiben eher ihr Vorgehen.

Eine Probandin betont, dass sie wenig Lust habe, Texte überhaupt zu korrigieren und mehr das Feedbackgespräch zu dem Schüler suchen und mit Fördermethoden arbeiten würde. Zwei weitere Gymnasiallehrkräfte betonen, dass sie Fazitkommentare unter dem Text verfassen. Fünf Lehrkräfte geben bei der Textbeurteilung direkte Verbesserungsvorschläge an. Im Gegensatz dazu gibt es einen Probanden, der lediglich Fehler unterstreicht und die Schüler*innen selber herausfinden lassen möchte, welches Defizit vorliegt. Mehrere Lehrkräfte betonen, dass ihre Beurteilungsmethoden von den jeweiligen Schüler*innen abhängen. Dabei wird ihr Hintergrundwissen zum Leistungsniveau des Schülers eingesetzt: *„Je nachdem, welches Hintergrundwissen ich habe. Also wenn das zum Beispiel Lese-Rechtschreibschwäche ist, dann mache ich da schon, dass ich dann dahin schreibe, wie es richtig ist, aber ich fordere das dann auch ein, dass die mir das verbessert zurückschicken“* (wNReaU2). Zudem spielt die Klassenstufe und die Relevanz der Beurteilung ebenso eine Rolle bei der Auswahl der Methodik: *„Wenn das jetzt eine Klausur in der Oberstufe wäre [...] dann muss man natürlich jetzt immer am Rand hier aufschreiben, das ist ein Rechtschreibfehler“* (mNGymUO2). Dieser Proband deutet hier auf die Korrekturrichtlinien zur Vorbereitung auf das Abitur hin. Andere Lehrkräfte betonen zudem, dass sie Klassenarbeiten präziser beurteilen als Texte, die im Unterricht geschrieben werden. Insgesamt haben die Lehrkräfte kaum bis gar keine Erklärungen und Begründungen für ihr Vorgehen genannt.

Diskussion

Im Nachfolgenden sollen die Ergebnisse zu den drei Forschungsfragen reflektiert und interpretiert werden. Dazu werden zunächst die Gemeinsamkeiten der Proband*innen in ihrem Textbeurteilungsverhalten und anschließend die Unterschiede aufgedeckt.

Gemeinsamkeiten der Geographielehrkräfte bei der Textbeurteilung

Die Analyse der Anforderung beim argumentativen Schreiben und dessen Beurteilung hat gezeigt, dass es sowohl für Lernende als auch für Lehrende eine herausfordernde Handlung ist (Zohar & Nemet, 2002; Kelly & Bazerman, 2003; Sandoval & Millwood, 2005; Spiegel 2006; Sampson & Clark 2008; Budke

& Uhlenwinkel, 2011; Petersen 2013; Becker-Mrotzeck, 2017; Stimming 2019). Lehrkräfte benötigen vielerlei Kompetenzen, um Schüler*innentexte zu beurteilen. Für die Studie wurden aus diesem Grund die Begrifflichkeiten der ‚Korrektur‘, ‚Textbeurteilung‘, ‚Diagnose‘ und ‚Feedback‘ definitorisch getrennt. In der Praxis der Lehrkräfte jedoch, ist ersichtlich geworden, dass diese Handlungen parallel und nicht chronologisch aufeinander aufbauend verlaufen und eine Trennung der Begrifflichkeiten nur implizit verläuft. Sie sehen diese Handlung gleichzeitig als Rückmeldung bzw. Feedback für die Schüler*innen an. Dies hat zum Vorteil, dass die Lehrkräfte eine effektive Zeitersparnis bei den Textbeurteilungen haben, kann jedoch auch dazu führen, dass die Inhalte und qualitativen Textbeurteilungskriterien nicht ausreichend berücksichtigt werden. Zudem könnte das Verständnis, dass Korrektur = Textbeurteilung = Feedback ist, dazu führen, dass die Schüler*innen einen sehr defizitorientierten Eindruck von ihrem Text bekommen, indem die Lehrkräfte sich vorwiegend an sprachlich textimmanente Fehlerkorrekturen bedienen. Zudem ist Feedback ein äußerst relevanter Teil des Lernens (Hattie, 2012) und sollte durch die Vorarbeit kriterienbasierter Diagnose, Textbeurteilung und Korrektur (Wieser, 2014; Müller & Ditton, 2014, Heid, 2015; Becker-Mrotzeck, 2020) separat und ausführlich geschehen.

Die Lehrkräfte haben teilweise betont, dass Textbeurteilung motivierend sei (siehe auch Beer, 2006) und nicht allein auf die Fehler und die Textimmanente konzentrieren sollte. Dies deckt sich ebenso mit den theoretischen Überlegungen von Becker-Mrotzeck (2020), welcher betonte, dass der Dialog zwischen Lehrkraft und Schüler*in bestenfalls durch Randkommentare und Textbeurteilung eröffnet werden sollte. Dementsprechend ist es zusammenfassend relevant, dass Lehrkräfte für die Unterschiede zwischen Korrektur, Textbeurteilung etc. sensibilisiert werden.

Die Ergebnisse aus den Interviews zeigen, dass die meisten Lehrkräfte angeben, mit Erwartungshorizonten und somit Kriterien zur Beurteilung von Schüler*innentexten zu arbeiten, was zunächst positiv ist, da dies eine genauere Beschreibung der Fehlerquellen und damit eine präzise Kompetenzdiagnose ermöglicht (Kleppin, 2007). Allerdings werden Erwartungshorizonte von den Lehrkräften in erster Linie bei Klausuren erstellt. Dies deckt sich mit Ergebnissen einer Studie, welche gezeigt hat, dass das Schreiben von Texten im Geographieunterricht für sie vor allem die Funktion der Leistungsbewertung hat (Steingrübl & Budke, 2022). Eine andere Überlegung ist, dass die Lehrkräfte Erwartungshorizonte bzw. Korrekturhilfsmittel als Last ansehen, viel Arbeit damit verbinden oder nicht erprobt mit den Hilfsmitteln sind und sie deswegen nur in Klausuren hinzuziehen.

Die Lehrkräfte korrigieren und beurteilen den Schülertext anhand impliziter Normen. Dies zeigt sich daran, dass sie diese oftmals nicht konkret formulieren, jedoch beurteilen können, ob es sich um einen gelungenen oder weniger gelungenen Text(abschnitt) handelt. Ebenso wenig Bewusstsein besteht beim Beurteilungsvorgang selbst: Es wurde festgestellt, dass nahezu alle Lehrkräfte Differenzen zwischen der vorherigen Beschreibung, wie sich ihr Korrekturvorgang gestaltet, und der tatsächlichen Handlung vorweisen. Zudem wurde bemerkt, dass die Lehrkräfte ihr Beurteilungshandeln kaum begründen.

Die Ergebnisse zeigen, dass nur wenige Lehrkräfte auf Materialbezug (Feilke et al., 2016; Schüler, 2017), Fachvokabular (Budke, 2021:3), räumliche und zeitliche Bedingungen (DGfG, 2020; Budke & Kuckuck, 2017), faktisch-korrekte Belege (Feser & Höttecke, 2016; Toulmin, 1996; Stapleton & Wu, 2015; Zohar & Nemet, 2002) und weitere fachlich-qualitativen Gütekriterien (Eignung, Relevanz, Gültigkeit)

achten. Diese geographiespezifischen Kriterien sind jedoch für die Bewertung der inhaltlichen Qualität der Argumentation entscheidend (Stapleton & Wu, 2015; Zohar & Nemet, 2002; Budke & Kuckuck, 2017; KMK, 2005; Tajmel, 2011; Busch & Ralle, 2013; Weiss et al. 2013). Dieses Erkenntnis vermittelt den Eindruck, dass die Geographielehrkräfte wenig Wert auf die fachliche Richtigkeit der Argumentation gelegt haben, obwohl sich insbesondere die fachliche Komponente von anderen Fächern unterscheidet und fokussiert werden sollte. Ein Erklärungsansatz könnte sein, dass sie nicht über spezielle Textbeurteilungsstrategien für argumentative Texte im Geographieunterricht verfügen und sie auf Richtlinien und Anforderungen aus ihrem Zweifach (was bei den Proband*innen dieser Studie oftmals eine Sprache war) zurückgreifen. Dies könnte ebenso erklären, warum die Beurteilung der sprachlichen und strukturellen Aspekte deutlich leichter fiel und häufiger durchgeführt wurde. Auch der aktuelle Forschungsstand aus den Fachdidaktiken zeigt, dass häufig ein sprachlicher Fokus bei der Textbeurteilung vorliegt (Rüede & Weber, 2009; Busch & Ralle, 2013; Weiss et. al., 2013; Feser & Höttecke, 2016). Dies spiegelt sich auch in der vorliegenden Studie wider. Dies sollte kritisch betrachtet werden, da sich das Fach Geographie ebenso wie andere Fachdidaktiken die Vermittlung und Intensivierung von Fachinhalten (Budke & Meyer, 2015) durch das Schreiben argumentativer Texte verspricht und sich daher der Fokus bei der Beurteilung der Texte ändern sollte. Dabei wäre es sinnvoll die Abiturrichtlinien anzupassen (Weiss et. al., 2013:12), die Beurteilungskriterien deutlicher zu differenzieren und angehende Lehrkräfte bereits in der Hochschule zu einer fachspezifischen Sichtweise und Beurteilung von argumentativen Texten zu sensibilisieren und auszubilden.

Unterschiede zwischen den Geographielehrkräften bei der Textbeurteilung

Die Lehrkräfte berichten in den Interviews, dass sie ihre Beurteilungsfähigkeiten durch „learning-by-doing“ entwickelt haben und keine konkreten Vorgaben hatten. Werden die Ergebnisse der Studie betrachtet, kann dies ein möglicher Grund dafür sein, dass die Lehrkräfte sehr individuell beurteilen und keine Standardisierung erkennbar ist – sie arbeiten mit unterschiedlichen Kürzeln, Methoden und Anforderungen an einen Text.

Dies zeigte sich auch in der Beurteilungszeit und Herangehensweise von holistischen Textbeurteilung bis hin zu sehr detaillierter Textkorrektur. Die Einstellungen zur Funktion der Textbeurteilung unterscheiden sich ebenfalls: Manche Lehrkräfte sind der Meinung, dass viele Fehlermarkierungen im Schüler*innentext die Schüler*innen demotivieren könnten und bedienen sich einer eher globalen Betrachtung des Textes und streichen weniger Fehler an. Andere Lehrkräfte sind sich einig, dass Vollständigkeit innerhalb der Korrektur relevant sei, damit die Schüler*innen alle Fehler erkennen und verbessern können. Darüber hinaus hinterfragen einige Lehrkräfte, inwiefern die Textkorrektur überhaupt Sinn macht und äußern Zweifel, dass die Schüler*innen die Korrekturen verarbeiten.

Diese Aspekte könnten dazu führen, dass die Schüler*innen, nie gänzlich nachvollziehen können, wie nun ein gelungener argumentativer Text aussieht und es ihnen schwerfallen könnte, eine Schreib- oder Berichtigungsroutine zu entwickeln. Daraus resultiert, dass eine gewisse Einheitlichkeit bei der Beurteilung von Vorteil sein kein. Lehrkräfte könnten somit in diesem Bereich durch Vorlagen und Beurteilungsraster sowie in ihrer didaktischen Einbettung der Leistungstransparenz unterstützt

werden. Dies scheint auch in Anbetracht der wechselnden Lehrkräfte der Schüler*innen während ihrer Schulzeit sinnvoll.

Zusammenfassend ist es wünschenswert, dass Schüler*innen vielseitiges Feedback zu argumentativen, strukturellen und sprachlichen Fehlern bekommen, damit sie ihre Argumentationskompetenzen weiter ausbauen können. Dazu kann es sinnvoll sein, die Förderung von Argumentations- und Korrekturkompetenzen der Geographielehrkräfte bereits in den Hochschulalltag zu integrieren, damit sie selbst argumentieren können, um dies anschließend vermitteln und diagnostizieren zu können. Dazu könnten Beurteilungs- und Korrekturraster zur Verfügung gestellt, entwickelt und erprobt werden.

Methodendiskussion/Limitation der Studie

Bei der Studie handelt es um ein unnatürliches Setting, welches unabhängig vom Schulalltag stattfand. Dadurch merkte ein Proband zur Korrektur an: „*Ich mach es gerade, glaub ich ausführlicher, als ich es tatsächlich im Unterricht machen würde*“ (mTGesUO). Dennoch bleibt an dieser Stelle gesagt, dass die Forschungsfragen gültig beantwortet werden konnten, da sich das Ziel darauf beschränkte, einen ersten Eindruck über Korrekturverfahren und Kriterien, die Geographielehrkräfte beim argumentativen Schreiben berücksichtigen, zu erhalten. Somit können in dieser qualitativen Forschung auch keine Generalaussagen über das Korrekturverhalten im Schulalltag getroffen werden.

Die Lehrkräfte haben den Schüler*innentext digital korrigiert, jedoch wurde teils von Proband*innen erwähnt, dass sie handschriftlich methodisch anders korrigiert hätten (Bsp. Unterschlängeln von Strukturfehlern, Zeichnen von Pfeilen). Zudem hängt die Art der Korrektur ebenso von dem Textbeispiel ab. Eventuell haben die Lehrkräfte auch weniger inhaltliche Kriterien korrigiert, da sie die Ausgangsmaterialien nicht kannten. Dennoch ändert sich die Relevanz der einzelnen Kriterien für einen gelungenen argumentativen Text nicht, sondern lediglich die Anzahl pro Kriterium.

Ein Problem bei der Nutzung der Korrekturraster (vgl. Tabelle 2; 3; 4) könnte sein, dass sich die einzelnen Kategorien der Korrekturraster teils stark bedingen und es somit zur Doppelkodierung kommen kann. Wenn eine Lehrkraft beispielsweise kritisiert, dass sie sich eine allgemeinere Einführung in den Konflikt im Text in der Einleitung gewünscht hätte, kann dies zum einen als mangelnde Darstellung der Strittigkeit und zum anderen auch als nicht-ausreichende Einleitung im strukturellen Sinn gewertet werden. Bezogen auf die Korrektur eines Schüler*innentextes kann es deshalb auch vorkommen, dass nicht immer trennscharf bewertet werden kann, ob es sich um sprachliche oder inhaltliche Fehlern handelt. Dies haben bereits andere Studien gezeigt (Budke & Weiss, 2014:14).

Conclusio

Die Studie hat gezeigt, dass Lehrkräfte den Schülertext nicht allumfassend bezüglich der drei Oberkategorien Sprache, Argumentation und Struktur beurteilten und Schwierigkeiten dabei aufweisen, konkrete Defizite vor allem im Bereich der Argumentation zu benennen und am Text zu korrigieren oder zu beurteilen. Die Zusammenstellung der Kriterien in den Tabellen zwei, drei und vier könnte nicht allein als Analysetool für das Forschungsdesign verwendet werden, sondern Geographielehrkräften auch Unterstützung bei der Textbeurteilung argumentativer Texte bieten.

Dabei sollte das Beurteilungsraster zunächst dazu genutzt werden, Stärken und Defizite eines Textes zu beschreiben und anschließend die Schüler*innenkompetenzen zu diagnostizieren, um Fördermethoden abzuleiten. Zudem könnte das Tool bei der Erstellung von Erwartungshorizonten eingesetzt werden. Wünschenswert für Lehrkräfte und Schüler*innen wäre es, wenn die Kriterien bewusst in den Unterricht integriert und transparent gemacht würden. Das könnte bei Schüler*innen zur Leistungstransparenz und Orientierung bei der Textproduktion beitragen. In einer weiteren Untersuchung könnte erforscht werden, inwiefern Schüler*innen überhaupt mit den Beurteilungen der Lehrkräfte umgehen. Dies ist von hoher Relevanz, um sowohl einen Lerneffekt bei den Schüler*innen zu erzeugen, als auch den Lehrkräften Sicherheit zu geben, dass die Schüler*innen ihre Beurteilungen nachvollziehen und nutzen können.

Abschließend kann festgehalten werden, dass Lehrkräfte Textbeurteilungskompetenzen in ihrem Schulalltag benötigen, um Auskunft über den Lernstand zu bekommen und den Schüler*innen Rückmeldung über ihre Leistung zu geben. Um Textbeurteilungen effektiv nutzen zu können, bedarf es jedoch zusätzlich an Diagnosefähigkeiten der Lehrkräfte bezüglich der Schüler*innenkompetenzen. Damit herrscht Forschungsbedarf im Bereich der Diagnosefähigkeiten der Lehrkräfte und inwiefern sie ihre Beurteilungen für die Diagnose der Schüler*innenkompetenzen nutzen, anschließend Rückmeldung darüber geben und dies im Unterricht einbetten.

Literatur

Abdollahzadeh, E., Amini Farsani, and M. Beikmohammadi (2017). Argumentative Writing Behavior of Graduate EFL Learners. *Argumentation*, 31(4), 641-661.

Allen, L. K., Likens, A. D., and McNamara, D. S. (2018). A multidimensional analysis of writing flexibility in an automated writing evaluation system. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. ACM.

Andrews, C. (2016). *Qualitative Argumentationsanalyse als Methode der empirischen Sozialforschung*. Dissertation. Bochum: Ruhr-Universität.

Aufschnaiter, C. von, Erduran, S., Osborne, J., and Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, 45(1), 101-131.

Bartosch, I., & Mayr, J. (2009). Wer beurteilen will, braucht Maßstäbe. In B. Hanfstingl, K. Krainer, & S. Zehetmeier (Hrsg.), *Fragen zur Schule – Antworten aus Theorie und Praxis* (S. 67-83). Studienverlag, Innsbruck.

Basel, N., Harms, U., and Pechtl, H. (2013). Analysis of students' arguments on evolutionary theory. *Journal of Biological Education*, 47(4), 192-199.

Basten, M., Kraft, A., and Wilde, M. (2017). Die Bedeutung der kommunikativen Einbettung für das Bewerten und schriftliche Argumentieren im Biologieunterricht. *Bildung und Erziehung*, 70(1), 57-74.

Benetos, K., and Betrancourt, M. (2020). Digital authoring support for argumentative writing: what does it change? *Journal of Writing Research*, 12(vol. 12 issue 1), 263-290.

Beer, R. (2006). Standards und Leistungsbeurteilung. Bedeutung und grundlegende Funktionen. *Information zur Deutschdidaktik, Zeitschrift für den Deutschunterricht in Wissenschaft und Schule*, 4, 52-63.

Becker-Mrotzek, M., Grabowski, J., and Steinhoff, T. (Hrsg.). (2017). *Forschungshandbuch empirische Schreibdidaktik*. Münster; New York: Waxmann.

Becker-Mrotzcek, M. & Böttcher, I. (2020): *Schreibkompetenz entwickeln und beurteilen*. 9. Aufl. Cornelsen: Berlin.

Böttcher, F., and Meisert, A. (2011). Argumentation in Science Education: A Model-based Framework. *Science & Education*, 20(2), 103-140.

Budke, A., Schiefele, U., and Uhlenwinkel, A. (2010). „I think it's stupid“ is no argument – some insights on how students argue in writing. *Teaching Geography*, 35(2), 66-69.

Budke, A., and Uhlenwinkel, A. (2011). Argumentieren im Geographieunterricht - Theoretische Grundlagen und unterrichtspraktische Umsetzungen. In C. Meyer, R. Henry, & G. Stöber (Hrsg.), *Geographische Bildung. Kompetenzen in der didaktischer Forschung und Schulpraxis*. Braunschweig.

Budke, A., and Weiss, G. (2014). Sprachsensibler Geographieunterricht. In M. Michalak (Hrsg.), *Sprache als Lernmedium im Fachunterricht. Theorien und Modelle für das sprachbewusste Lehren und Lernen*. Baltmannsweiler, 113-133.

Budke, A., and Meyer, M. (2015). Fachlich argumentieren lernen – Die Bedeutung der Argumentation in den unterschiedlichen Schulfächern. In A. Budke, M. Kuckuck, M. Meyer, F. Schäbitz, K. Schlüter, & G. Weiss (Hrsg.), *Fachlich argumentieren lernen. Didaktische Forschungen zur Argumentation in den Unterrichtsfächern* (9-30). Münster; New York: Waxmann.

Budke, A., & Kuckuck, M. (Hrsg.) (2017). Sprache im Geographieunterricht. In: Sprache im Geographieunterricht. Bilinguale und sprachensible Materialien und Methoden. 7-38. Münster; New York: Waxmann.

Budke, A., Gebele, D., Königs, P., Schwerdtfeger, S., & Zepter, A. (2020a). Student texts produced in the context of material-based argumentative writing: Interdisciplinary research-related conception of an evaluation tool. RISTAL, 3, 108-125. Aufgerufen am 27.07.2023 von https://www.ristal.org/fileadmin/user_upload/RISTAL/PDFs_2020/RISTAL_Budke_DOI.pdf.

Budke, A. and Schäbitz, F. (2021, Hrsg.): Argumentieren und Vergleichen. Beiträge aus der Perspektive verschiedener Fachdidaktiken. Berlin.

Bitchener, J., and Knoch, U. (2008). The value of written corrective feedback for migrant and international students. Language Teaching Research, 12(3), 409-431. <https://doi.org/10.1177/1362168808089924>

Busch, H., & Ralle, B. (2013). Diagnostik und Förderung fachsprachlicher Kompetenzen im Chemieunterricht. In M. Becker-Mrotzek, K. Schramm, E. Thürmann, & H. J. Vollmer (Hrsg.), Sprache im Fach. Sprachlichkeit und fachliches Lernen (277-294). Münster; New York: Waxmann.

Chase, B. J. (2011). An Analysis of the Argumentative Writing Skills of Academically Underprepared College Students. Dissertation. Chase_columbia_0054D_10083.pdf.

Crystal, D. (2008). A Dictionary of Linguistics and Phonetics (6th ed.). Blackwell Publishing.

Deutsche Gesellschaft für Geographie (DGFG) (Hrsg.) (2020): Bildungsstandards im Fach Geographie für den mittleren Schulabschluss mit Aufgabenbeispielen. 8., aktualisierte Auflage. Bonn: Deutsche Gesellschaft für Geographie, Aufgerufen am 27.10.2022 von https://geographie.de/wp-content/uploads/2020/09/Bildungsstandards_Geographie_2020_Web.pdf.

Dimbath, O., Ernst-Heidenreich, M., & Roche, M. (2018). Praxis und Theorie des Theoretical Sampling. Methodologische Überlegungen zum Verfahren einer verlaufsorientierten Fallauswahl. FQS 19(3), Art. 34.

Ditton, H., & Müller, A. (2014). Feedback: Begriff, Formen und Funktionen. In H. Ditton & A. Müller (Hg.), Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder. 11-28. Münster; New York: Waxmann.

Dudenredaktion (o.J.): „Korrektur“ auf Duden Online. Aufgerufen am 11.07.2023 von <https://www.duden.de/rechtschreibung/Korrektur>.

Ellis, R. (2009). Corrective Feedback and Teacher Development. L2 Journal, 1(1), 3-18.

Erduran, S., and Jiménez-Aleixandre, M.P. (2007). Argumentation in Science Education: an Overview. In Argumentation in Science Education: Perspectives from Classroom-Based Research. S. Erduran., M. P. Jiménez- Aleixandre (Hrsg.). 3-27. Springer.

Feilke, H., & Bachmann, T. (2014). Werkzeuge des Schreibens. Beiträge zu einer Didaktik der Textprozeduren. Stuttgart: Filibach bei Klett.

Feilke, H., Lehnen, K., Rezat, S., & Steinmetz, M. (2016). Materialgestütztes Schreiben lernen. Grundlagen. Aufgaben, Materialien. Braunschweig: Westermann Schroedel.

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), Feedback in Second Language Writing: Contexts and Issues (Cambridge Applied Linguistics, pp. 81-104). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139524742.007

Feser, M. S., & Höttecke, D. (2016). Wie Physiklehrkräfte Schülertexte beurteilen – Instrumententwicklung. In C. Maurer (Hrsg.), *Implementation fachdidaktischer Innovation im Spiegel von Forschung und Praxis* (123-126). Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Zürich 2016.

Feser, M. S., Höttecke, D., & Ehmke, T. (2016). Testitems zur qualitativen Untersuchung der Ressourcen von Physiklehrkräften beim Bewerten schriftlicher Schülerleistungen. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung*. Aufgerufen am 11.07.2022 von <http://www.phydid.de/index.php/phydid-b/article/view/677/819>.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New Brunswick: Aldine.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

Heid, H. (2015). Über Relevanz und Funktion des Fehlerkriteriums. In Gartmeier et al. (Hrsg.), *Funktionen von Fehlern im Kontext individueller und gesellschaftlicher Entwicklung*. 33-52. Münster; New York:: Waxmann.

Heine, L., & Schramm, K. (2016). Introspektion. In K. Schramm, K. Schramm, & F. Klippel, M. K. Legutke (Eds.), *Forschungsmethoden in der Fremdsprachendidaktik: Ein Handbuch* (173-181). Narr Francke Attempto Verlag.

Helmke, A. (2012). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Kallmeyer und Klett.

Hindmarsh, K. & Budke, A. (2023): What are the competences & difficulties of school students in writing material-based & multi-perspective argumentations in geography classrooms? In: *European Journal of Geography*. Vol.14, Issue 3. S. 55-67. DOI: <https://doi.org/10.48088/ejg.k.hin.14.3.055.067>

Jürgens, E. (2010). *Leistung und Beurteilung in der Schule – Eine Einführung in Leistungs- und Bewertungsfragen aus pädagogischer Sicht* (7. Auflage). Academia Verlag.

Kleppin, K. (2007). Fehler und Fehlerkorrektur. In *Fernstudienprojekt zur Fort- und Weiterbildung im Bereich Germanistik und Deutsch als Fremdsprache – Teilbereich Deutsch als Fremdsprache*; 19 (1-11). Berlin: Langenscheidt.

Knudson, R. E. (1992). Analysis of Argumentative Writing at Two Grade Levels. *The Journal of Educational Research*, 85(3), 169-179.

Kultusministerkonferenz (KMK). (2005). *Einheitliche Prüfungsanforderungen in der Abiturprüfung Geographie*. Aufgerufen am 09.07.2023 von https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Geographie.pdf.

Konrad, K. (2020). Lautes Denken. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie* (1-12). Springer.

Kopperschmidt, J. (2000). *Argumentationstheorie*. Hamburg: Junius.

Lam, Y. W., Hew, K. F., & Chiu, K. F. (2018). Improving argumentative writing: Effects of a blended learning approach and gamification. *Language Learning and Technology*, 22(1), 97-118.

Ditton, H., and Müller, A. (Hrsg.). (2014). *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder*. Münster; New York: Waxmann.

Feilke, H., and Pohl, T. (Hrsg.). (2014). *Deutschunterricht in Theorie und Praxis. (DTP). Handbuch zur Didaktik der deutschen Sprache und Literatur in elf Bänden. Band 4: Schriftlicher Sprachgebrauch - Texte verfassen*. Schneider Hohengehren.

- Petersen, I., (2013): Entwicklung schriftlicher Argumentationskompetenz bei ein- und mehrsprachigen Oberstufenschüler/-innen und Studierenden. In: Brandl, H., Langelahn, E., Riemer, C. (Hrsg.) (2013): Mehrsprachig in Wissenschaft und Gesellschaft. Mehrsprachigkeit, Bildungsbeteiligung und Potenziale von Studierenden mit Migrationshintergrund. Bielefeld. 69-79.
- Quasthoff, U., & Domenech, M. (2016). Theoriegeleitete Entwicklung und Überprüfung eines Verfahrens zur Erfassung von Textqualität (TexQu) am Beispiel argumentativer Briefe in der Sekundarstufe I. *Didaktik Deutsch*, 21(41), 21-43.
- Rae, A. M., & Cochrane, D. K. (2008). Listening to students. How to make written assessment feedback useful. *Active Learning in Higher Education*, 9(3), 217-230.
- Riemeier, T., von Aufschnaiter, C., Fleischhauer, J., & Rogge, C. (2012). Argumentationen von Schülern prozessbasiert analysieren: Ansatz, Vorgehen, Befunde und Implikationen. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 181-200.
- Rüede, C., & Weber, C. (2009). Keine Diagnose ohne Auseinandersetzung mit Form, Inhalt und Hintergrund von Schülertexten. In *Beiträge zum Mathematikunterricht 2009* (819-822).
- Sampson, V., & Clark, D. (2008). Assessment of the Ways Students Generate Arguments in Science Education: Current Perspectives and Recommendations for Future Directions. *Science Education*, 92(3), 447-472.
- Sandoval, W. A., & Millwood, K. A. (2005). The Quality of Students' Use of Evidence in Written Scientific Explanations. *Cognition and Instruction*, 23(1), 23-55.
- Schüler, L. (2017). Materialgestütztes Schreiben argumentativer Texte. Untersuchungen zu einem neuen wissenschaftspropädeutischen Aufgabentyp in der Oberstufe. In B. Rothstein (Hrsg.), *Thema Sprache – Wissenschaft für den Unterricht* (Bd. 25, 1-23). Baltmannsweiler: Schneider Verlag Hohengehren GmbH.
- Simon, S., Erduran, S., & Osborne, J. (2006). Learning to Teach Argumentation: Research and development in the science classroom. *International Journal of Science Education*, 28(2-3), 235-260.
- Spiegel, C., (2006): Argumentieren lernen im Unterricht- ein funktional didaktischer Ansatz. In: Grundler, E., R. Vogt (Hrsg.) (2006): *Argumentieren in Schule und Hochschule. Interdisziplinäre Studien.* Staufenburg Verlag, Tübingen. 63- 76.
- Stapleton, P., & Wu, Y. A. (2015). Assessing the quality of arguments in students' persuasive writing: A case study analyzing the relationship between surface structure and substance. *Journal of English for Academic Purposes*, 17, 12-23.
- Steingrübl, S., & Budke, A. (2022). Writing in Geography Lessons—An Unreflected Routine? *Education sciences*, 12(587), 1-23.
- Stimming, E., (2019): Argumentieren im fächerverbindenden Deutsch- und Sachunterricht- erste Ergebnisse und Schlussfolgerungen aus einer Interviewstudie mit Lehrkräften. In: Becker- Mrotzek, M., J. Jost, T. Pohl, K. Schindler (2019): *Kölner Beiträge zur Sprachdidaktik. Reihe A. Gilles und Francke Verlag. Duisburg.* 31- 48.
- Tajmel, T. (2011). Wortschatzarbeit im mathematisch-naturwissenschaftlichen Unterricht. *ide*, 35(1), 83-93.
- Toulmin, S. (1996). *Der Gebrauch von Argumenten* (2. Aufl.). Weinheim.
- van Ophuysen, S., and Lintorf, K. (2013). Pädagogische Diagnostik im Schulalltag. In A. Herausgeber & B. Herausgeber (Hrsg.), *Lernen in Vielfalt. Chance und Herausforderung für Schul- und Unterrichtsentwicklung.* 55-76. Münster; New York: Waxmann.

Weber, C., & Ruede, C. (2010). Schülertexte aus unterschiedlichen diagnostischen Perspektiven lesen. In A. Lindmeier & S. Ufer (Hrsg.), Beiträge zum Mathematikunterricht 2010 (919–922). Munster: WTM.

Weiss, G., Gorbandt, E., Mäsgen, J. & Wiktorin, D. (2013). Zwischen Materialschlacht und Reproduktion - Schriftliche Zentralabituraufgaben im Bundesländervergleich und erste Erkenntnisse zu Schülerleistungen. Geographie und Schule, H.206, 4-14.

Wieser, A. (2014). Leistungsbeurteilung im kompetenzorientierten Geographie- und Wirtschaftskundeunterricht. Abgerufen am 28.09.2023, von <https://unipub.uni-graz.at/obvugrhs/download/pdf/240212?originalFilename=true>.

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 39(1), 35-62.