

Martin Nitsche, Monika Waldis (Hrsg.)

Historisches Denken erforschen

gh
15

Geschichtsdidaktik heute





Geschichtsdidaktik heute
www.geschichtsdidaktik-heute.ch

In der Reihe «Geschichtsdidaktik heute» werden Arbeiten veröffentlicht, die sich mit innovativen Ansätzen und methodischer Klarheit der empirischen Erforschung theoretischer Konzepte sowie der Rückbindung empirischer Befunde an die Theoriebildung im Feld der Geschichtsdidaktik widmen.

Herausgeberschaft der Reihe «Geschichtsdidaktik heute» im hep Verlag

Prof. em. Dr. Béatrice Ziegler
Politische Bildung und Geschichtsdidaktik, Pädagogische Hochschule FHNW
am Zentrum für Demokratie Aarau

Prof. em. Dr. Markus Furrer
Em. Professor für Geschichte und Geschichtsdidaktik an der
Pädagogischen Hochschule Luzern

Stephan Hediger
Bereichsleiter Geografie, Geschichte, Religion und Kultur,
Pädagogische Hochschule Zürich

Prof. Dr. Nadine Ritzer
Professorin am Zentrum für Educational Governance und
Demokratiebildung der Pädagogischen Hochschule Zürich

Martin Nitsche und Monika Waldis (Hrsg.)

Historisches Denken erforschen

Reihe «Geschichtsdidaktik heute», Band 15





*Publiziert mit Unterstützung des Schweizerischen Nationalfonds
zur Förderung der wissenschaftlichen Forschung.*

Publiziert von
hep Verlag AG
Gutenbergstrasse 31
3011 Bern
hep-verlag.com

ISBN Print: 978-3-0355-2658-5

DOI: <https://doi.org/10.36933/9783035526585>



Dieses Werk ist lizenziert unter einer Creative Commons
Namensnennung 4.0 International Lizenz.

Inhaltsverzeichnis

Einführung	7
Historisches Denken erforschen	9
<i>Martin Nitsche und Monika Waldis</i>	
Assessment historischen Denkens	29
Historische Kompetenzen im Umgang mit Medienberichten über den 6. Januar 2021	31
<i>Paul G. Fitchett, Carolin Hestler und Franziska Rein</i>	
Das Repertory-Grid-Interview als qualitativ empirische Forschungsmethode in der Geschichtsdidaktik	51
<i>Franziska Rein</i>	
Den Umgang mit Quellen und Darstellungen digital beforschen	72
<i>Jan Scheller, Jonas Schobinger, Martin Nitsche</i>	
Digitalisierung und historisches Denken	97
Künstliche Intelligenz im Fach Geschichte	99
<i>Christiane Bertram, Zarah Weiß, Lisa Zachrich, Ramon Ziai</i>	
«In ihrer Zeitung [...] versuchten sie an alle heranzukommen und sie an ihre tolle Kultur und an sie selbst zu erinnern»	124
<i>Kristopher Muckel</i>	

Sprachsensible Erforschung historischen Denkens	147
Historisches Erklären lernen?	149
<i>Tobias Flink</i>	
Zur sprachbasierten Diagnose historischen Denkens in Texten von Schüler*innen	171
<i>Viola Schrader</i>	
Abstracts	193
Autor*innen	197
Gutachter*innen der Beiträge	200

«In ihrer Zeitung [...] versuchten sie an alle heranzukommen und sie an ihre tolle Kultur und an sie selbst zu erinnern»

(Digitalisierte) historische Zeitungen als Quellen im Geschichtsunterricht¹

Kristopher Muckel

1 Einleitung

Waren Zeitungen in ihrer Anfangszeit im 17. Jahrhundert vorrangig eine Zuverdienstmöglichkeit für Drucker in stark frequentierten Ortschaften, entwickelte sich insbesondere im Umfeld der Krisen und Konflikte des 18. Jahrhunderts ein reges Pressewesen (Stöber, 2005, 61–77), aus dem im 19. Jahrhundert eine Massenpresse entstand, deren Produkte mit Fug und Recht die Stellung von Leitmedien für sich beanspruchten. Konnte die gedruckte Zeitung diese Rolle in der Presselandschaft, die sich nach dem Zweiten Weltkrieg in den westdeutschen Besatzungszonen bzw. der Bundesrepublik Deutschland entwickelte, seit den 1950er-Jahren zum Teil wieder einnehmen (Bösch, 2019, 6–9), geriet die traditionelle Zeitung spätestens in den 2000er-Jahren mehr und mehr unter den Druck digitaler Angebote (Fleischhacker, 2014, 135–145).

Darf der (gedruckten) Zeitung damit in der Gegenwart – wenn überhaupt – nur noch begrenzt eine prägende Rolle für den gesellschaftlichen Diskurs zugeschrieben werden, ist ihre Bedeutung für die politische und gesellschaftliche Entwicklung mindestens im 19. und 20. Jahrhundert wohl unumstritten. Inwiefern daraus auch ein Potenzial für historisches Lernen in gegenwärtigem Geschichtsunterricht erwachsen kann, ist die zentrale Frage des Projektes *Historische Zeitungen und digitale Zeitungsarchive im Geschichtsunterricht*. In dessen Rahmen werden im Folgenden bisherige geschichtsdidaktische Auseinandersetzungen mit historischen Zeitun-

¹ Das Zitat im Titel ist eine Lernendenäußerung (Dokument SaErUm04 1) im Rahmen der unter 3 beschriebenen Studie.

gen diskutiert (2). Auf Grundlage dieser Vorarbeit wird die im Projekt durchgeführte empirische Studie vorgestellt (3) und ein Teilbereich der Auswertung im Detail präsentiert (4). Abschliessend werden deren Ergebnisse diskutiert, in den Kontext des Gesamtprojektes eingeordnet und weiterführende Fragen aufgeworfen (5).

2 Die historische Zeitung in der geschichtsdidaktischen Diskussion

Die geschichtsdidaktische Auseinandersetzung mit historischen Zeitungen konzentrierte sich bislang vor allem auf aus den definatorischen Eigenschaften von Zeitungen abgeleiteten Annahmen zu ihrer Relevanz für historisches Lernen sowie auf entsprechende Praxishinweise und Unterrichtsentwürfe (Kuchler, 2016, 34–39). Zur Beschreibung der Wesensmerkmale der Zeitung haben sich die von Groth vorgeschlagenen Merkmale der Periodizität, Universalität, Aktualität und Publizität allgemein durchgesetzt (Groth, 1960, 101–102). Sie bieten gerade aus gegenwärtiger Perspektive den Vorteil, Zeitungen nicht über ihre Materialität zu definieren, sodass ihr Anwendungsbereich auch auf digital publizierte Materialien erweitert werden kann, die den Anspruch verfolgen, regelmässig über alle von ihren Leser*innen als relevant betrachtete Themen aktuell zu berichten und (zumindest potenziell) frei zugänglich zu sein.

Insgesamt deutet die bisherige geschichtsdidaktische Auseinandersetzung mit historischen Zeitungen, die insbesondere von Michael Sauer geprägt worden ist (z. B. Sauer, 2006), darauf hin, dass diese im Geschichtsunterricht insbesondere im Hinblick auf die Vermittlung von alltags- und lokalgeschichtlichen Themen einerseits sowie andererseits in Bereichen mit recht deutlich divergierenden Perspektiven, die sich in verschiedenen Zeitungen niederschlagen können, nutzbringend für historisches Lernen eingesetzt werden können. Allerdings scheint dies bislang nicht zu passieren (Kuchler, 2016, 39–43), was umso mehr erstaunt, da historische Zeitungen bereits seit geraumer Zeit in enormer Zahl frei im Netz verfügbar sind.

In der Verknüpfung von Zeitungsdigitalisaten mit digitalen Suchumgebungen liegen, so die im Weiteren verfolgte These, Potenziale für das historische Lernen, die über den beschriebenen Nutzen der historischen Zeitung im Geschichtsunterricht hinausgehen. Es wird angenommen, dass sich diese aus den heuristischen²

2 Der Begriff der Heuristik wird bereits grundlegend definiert bei Droysen (1868, 15) sowie darauf aufbauend auch von Rüsen (2013, 171).

Anforderungen ergeben, die sich stellen, wenn Lernende eigenständig Zeitungen zur Beantwortung konkreter Fragestellungen recherchieren und auswählen. So verlangt eine Recherche, die im Sinne Bernsens (2017, 314) über die bloße Suche nach Sachinformationen hinausgeht, zunächst eine grundlegende Einsicht in die Suchlogik des jeweiligen Digitalarchivs, mithin ein Verständnis der Funktionsweise der Suchfunktionen und Navigationsmöglichkeiten im entsprechenden Hypertext, als welcher das Digitalarchiv anzusehen ist (Kramersitsch, 2007, 130–147). Diese fachunspezifisch anmutende Kompetenz wird dadurch Teil eines historischen Denkprozesses, dass sie auf den historischen Gegenstand bezogen wird (Hodel, 2016, 155), also auf zentrale Aspekte der historischen Quellenkritik³ zurückgreifen muss, um die Brauchbarkeit der aufgefundenen Suchergebnisse für die Fragestellung zu überprüfen.

3 Vorstellung Studiendesign und Auswertungsinstrument

Die Überprüfung dieser theoretischen Annahmen zum Wert historischer Zeitungen und digitaler Zeitungsarchive für das historische Lernen ist das wesentliche Erkenntnisinteresse des Projektes. Es ordnet sich damit in den Kontext der Medienvergleichsstudien mit geschichtsdidaktischem Erkenntnisinteresse ein, wobei zweischrittig vorgegangen wird und in der ersten Forschungsfrage zunächst der grundsätzliche Nutzen historischer Zeitungen für das historische Lernen im Mittelpunkt steht:

Können im Vergleich zwischen Lernenden, die mit didaktisierten Quellen gearbeitet haben (Gruppe Q), und Lernenden, die mit verschiedenen Formen historischer Zeitungen gearbeitet haben (Gruppe Z), Unterschiede in Schreibprodukten zu identischen Lernaufgaben identifiziert werden?

Die Offenheit im Hinblick auf den Zeitungseinsatz verfolgt das Ziel einer zunächst grundlegenden Identifikation möglicher Potenziale des Mediums für das historische Lernen. Auf die Unterscheidung der verschiedenen Einsatzformen zielt die zweite Frage ab:

Können im Vergleich zwischen Lernenden, die mit Faksimiledrucken historischer Zeitungsausgaben (Gruppe ZA), Lernenden, die mit digitalisierten Faksimiles der gleichen Zeitungen auf Tablets (Gruppe ZD), und Lernenden, die

3 Überlegungen zu einer Quellenkritik im Digitalen, deren Notwendigkeit z.B. Wineburg (2019, 117–118) betont, befinden sich, so der Befund bei Föhr (2017, 8), noch im Anfangsstadium.

selbstständig in Digitalarchiven nach historischen Zeitungen recherchiert haben (Gruppe DA), Unterschiede in Schreibprodukten zu identischen Lernaufgaben identifiziert werden?

Um Prozesse des historischen Lernens sicht- und messbar zu machen, wurden diese anhand ihrer Manifestationen in Lernendennarrationen zu Schreibaufgaben untersucht. Dabei wurde von der Prämisse ausgegangen, dass jede Interpretation der Vergangenheit grundsätzlich dadurch eine Narration darstellt, dass sie sich von der Gegenwart aus deutend Bereichen der Vergangenheit zuwendet (Ankersmit, 1994, 33, 40) und dazu erklärende und argumentierende Elemente integriert (Rüsen, 1986, 37–42). Die Beschreibung der Lernendentexte als Narrationen umfasst somit deren gesamte Spannweite an Sprachhandlungen, die durch die deutende Ausrichtung auf die Vergangenheit historisch werden.

Die Erhebung der Lernendentexte erfolgte im Rahmen einer quasi-experimentellen⁴ Interventionsstudie⁵ im Prä-Post-Design⁶ im ersten Halbjahr des Schuljahres 2020/2021. 63 Schüler*innen in vier Lerngruppen der Jahrgangsstufe 10 eines nordrhein-westfälischen Gymnasiums haben daran teilgenommen. In zwei Pilotierungen wurden im zweiten Halbjahr des vorausgegangenen Schuljahres Aufgabenstellungen für die Schreibaufgaben, die Auswertungsinstrumente, verschiedene Möglichkeiten der Darreichung der Zeitungen an die Lernenden (Druckformate, Dateiformate, Endgeräte) mit anderen Schüler*innen als jenen in der Haupterhebung erprobt. Eine ursprünglich geplante Randomisierung der Treatmentgruppen auf Grundlage einer zu Schuljahresbeginn durchzuführenden Fragebogenerhebung und der fachlichen Leistungen der beteiligten Lernenden im vorausgegangenen Schuljahr konnte aufgrund der schulbezogenen Einschränkungen der Coronapandemie nicht umgesetzt werden. Zudem musste der Erhebungszeitraum erheblich verkürzt und die methodische Ausgestaltung der Unterrichtseinheiten drastisch auf individuelle Lernformate, die keine Platzwechsel der Lernenden im Raum erforderlich machten, eingeschränkt werden. Auch wenn die Vergleichbarkeit der Erhebungsgruppen damit nicht auf dem vorgesehenen Niveau sichergestellt werden konnte, wird

4 Zur Charakterisierung quasiexperimenteller Settings vgl. z.B. Döring und Bortz (2016, 193–194).

5 Eine allgemeine Definition der Intervention als gezielten Eingriff in ein System bietet Leutner (2012, 18), zur Spannweite von Interventionsmassnahmen im pädagogischen Feld vgl. Souvignier und van Dignath Ewijk (2010, 14).

6 Den Einsatz von Prä-Post-Designs für den Nachweis kurzfristiger treatmentbedingter Veränderungen diskutiert Hagenauer (2010, 245).

die Aussagekraft der Hauptstudie durch den Fokus auf die individuellen relativen Veränderungen in der Bearbeitung der Schreibaufgaben zwischen den Erhebungszeitpunkten sowie die Kontrolle der Störeinflüsse durch flankierende Begleitstudien abgesichert.

In deren Rahmen wurden die Schüler*innen mittels eines halboffenen Fragebogens sowohl grundsätzlich im Hinblick auf ihre Einstellung zum Fach Geschichte als auch bezüglich ihrer allgemeinen und schulischen Erfahrung mit Zeitungen befragt. Zur Sicherung der Vergleichbarkeit des Unterrichts, der von den schulischen Lehrkräften auf Grundlage von durch den Verfasser erstellten Materialien durchgeführt worden ist, wurde der Unterricht systematisch beobachtet und protokolliert. Für die Haupterhebung ergab sich schlussendlich das in Abb. 1 dargestellte Design:

Hauptstudie (geplant Mitte August bis Dezember 2020, durchgeführt Oktober bis Dezember 2020)
 Vier Lerngruppen der JGS EF (10) eines Gymnasiums in NRW

<p>Vorlauf</p> <ul style="list-style-type: none"> • Thematische Einführung im Unterricht • Schulung der Lehrkräfte • Schulung der Beobachter*innen, Beginn der Unterrichtsbeobachtung • Klärung der Rahmenbedingungen 	<p>Prätest</p> <ul style="list-style-type: none"> • Erhebung der Trifftigkeit von Lernendentexten → materialgestützte Schreibaufgabe (zwei didaktisierte Quellen) • Erhebung von Störeinflüssen auf das Schreibergebnis → Halboffener Fragebogen SuS 	<p>Treatment</p> <ul style="list-style-type: none"> • Inhaltsgleiche und methodenvergleichbare, aber materialdifferenzierte Unterrichtserteilung durch die schulischen Lehrkräfte in den vier Erhebungsgruppen <ul style="list-style-type: none"> • digitale Zeitungsarchive • (ausgewählte) digitalisierte Zeitungsausgaben • (ausgewählte) Faksimiledrucke • didaktisierte Quellen • Systematische Unterrichtsbeobachtung 	<p>Posttest</p> <ul style="list-style-type: none"> • Erhebung der Trifftigkeit von Lernendentexten → materialgestützte Schreibaufgabe (auf Grundlage der Materialien des Treatments) • Erhebung von Störeinflüssen auf das Schreibergebnis → Halboffener Fragebogen SuS
<p>6 UE (45 Min.)</p>	<p>2,5 UE</p>	<p>10 UE</p>	<p>2,5 UE</p>

Abb. 1: Überblick Design Hauptstudie

Das Treatment selbst erfolgte in drei Schritten. So erhielten die Gruppen erstens eine Einführung in die Arbeit mit ihren jeweiligen Medien sowie eine grundlegende Schulung für die Verwendung multipler Dokumente für eine Schreibauf-

gabe (z. B. Goldman, 2004, 343–344).⁷ Die auf diesem Wege eingeführten Strategien zum Umgang mit historischen Zeitungen (DA/ZD/ZA), deren Digitalisaten (DA/ZD) sowie zur Recherche in Digitalarchiven (DA) wurden in einem zweiten Schritt unter Anleitung der Lehrkräfte ebenso trainiert wie jene zum Umgang mit den verschiedenen Formen didaktisierter Quellen, die in der entsprechenden Gruppe (Q) zum Einsatz gekommen sind. In diesem Kontext wurde drittens auch die inhaltliche Grundlage für die Schreibaufgabe des Posttests gelegt.

Im Folgenden wird ein Teilbereich der Auswertung der Lernendentexte im Fokus stehen. Eingebettet war die Erhebung in den Themenbereich «Fremdsein, Vielfalt und Integration – Migration am Beispiel des Ruhrgebiets im 19. und 20. Jahrhundert» des Inhaltsfeldes I des Kernlehrplans des Landes Nordrhein-Westfalen für die gymnasiale Oberstufe (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2014, 23). Da aufgrund des verkürzten Erhebungszeitraums die Einführung von für die Lernenden neuen Aufgabentypen nicht mehr realisierbar gewesen und von schulischer Seite eine Einbindung der Klausurvorbereitung in die Erhebungsphase eingefordert worden war, wurde für die Erhebungsaufgaben trotz der Kritik, die unter anderem Schönemann, Thünemann und Zülsdorf-Kersting (2011, 17–18) vorgebracht haben, auf Aufgabentypen im Sinne der Einheitlichen Prüfungsanforderungen für das Abitur zurückgegriffen (Kultusministerkonferenz, 2005, 10). Aufbauend auf einer Analyseaufgabe zu den nachfolgend genannten Quellen wurden für die beiden Testzeitpunkte die in Tabelle 1 zusammengefassten Aufgaben formuliert. Die so weit als möglich gleich gehaltenen Aufgabenstellungen sollten die entstandenen Schreibprodukte sowohl hinsichtlich ihres Umfangs als auch des Umgangs mit der Aufgabenstellung vergleichbar machen. Ein Gewöhnungseffekt wurde dafür vor dem Hintergrund in Kauf genommen, dass die erweiterte Materialgrundlage der zweiten Erhebung einen solchen abschwächt und eine bloße Reproduktion aus der Erinnerung an den ersten Erhebungszeitpunkt verhinderte.

7 Der Forschungsstand zum Einsatz multipler Dokumente als Grundlage für Schreibaufgaben kann hier nicht umfassend diskutiert werden, wird aber in der Gesamtstudie berücksichtigt.

Erster Erhebungszeitpunkt			
Material- grundlage	<ul style="list-style-type: none"> - Die Polen im rheinisch-westfälischen Steinkohlen-Bezirke, Auszug aus Gau «Ruhr und Lippe» des Alldeutschen Verbandes (1901, 55) - Übersetzungen aus dem Wiarus Polski, Nr. 1, Bochum, den 1. Januar 1901 (Institut für Zeitungsforschung, Dortmund Hz301) 		
Aufgaben- stellungen	<ol style="list-style-type: none"> 1. Erläutere zunächst kurz, was aus Sicht der Migrationsforschung unter Integration zu verstehen ist, und 2. beurteile auf dieser Grundlage, inwiefern die Bedingungen für eine solche im Ruhrgebiet gegeben waren (beziehe dich dabei explizit auf die beiden vorliegenden Quellen). 3. Bewerte abschliessend, inwiefern (aus heutiger Sicht) die Haltungen der sich selbst als deutsch und polnisch bezeichnenden Gruppierungen nachvollziehbar sind. Beziehe ggf. auch aktuelle Beispiele mit ein. 		
Zweiter Erhebungszeitpunkt			
Material- grundlage	Digitalarchive	Faksimiles historischer Zeitungen (analog bzw. als Digitalisat)	didaktisierte Quellen (vor- ausgewählte, gekürzte, kommentierte und trans- skribierte Quellen)
	<p>Recherche für den Ereigniszeitraum um 1905 in:</p> <p>Zeitungen aus dem Ruhrgebiet, z. B.</p> <ul style="list-style-type: none"> - Velberter Zeitung, - Dortmunder Zeitung; <p>Regionalzeitungen aus der Umgebung des Ruhrgebiets, z. B.</p> <ul style="list-style-type: none"> - Münsterischer Anzeiger (alle in https://zeitpunkt.nrw/); <p>reichsweit erscheinenden Zeitungen, z. B.</p> <ul style="list-style-type: none"> - Vossische Zeitung (in http://zefys.staatsbibliothek-berlin.de/) <p>Gewerkschaftszeitungen, z. B.</p> <ul style="list-style-type: none"> - Bergarbeiterzeitung (in: http://library.fes.de/gewerkschaftszeitung/) 	<p>Gesamtausgaben der:</p> <p>Zeitungen aus dem Ruhrgebiet:</p> <ul style="list-style-type: none"> - Velberter Zeitung (11. 01. 1905), - Dortmunder Zeitung (11. 01. 1905); <p>Regionalzeitung aus der Umgebung des Ruhrgebiets</p> <ul style="list-style-type: none"> - Münsterischer Anzeiger (11. 01. 1905) <p>reichsweit erscheinenden Zeitung</p> <ul style="list-style-type: none"> - Vossische Zeitung (12. 01. 1905 Morgenausgabe) <p>Gewerkschaftszeitung</p> <ul style="list-style-type: none"> - Bergarbeiterzeitung (14. 01. 1905) 	<ul style="list-style-type: none"> - gekürzter Quellenauszug aus Henning und Sieg (1997, 26–27) - drei gekürzte Quellenauszüge aus Tenfelde und Urban (2010, 483–485)

	Zweiter Erhebungszeitpunkt
Aufgabenstellungen	Arbeitsaufträge: <ol style="list-style-type: none">1. Erläutere zunächst noch einmal kurz, was aus Sicht der Migrationsforschung unter Integration zu verstehen ist, und2. beurteile auf dieser Grundlage, inwiefern eine solche mit Blick auf die polnischsprachigen Migranten im Ruhrgebiet stattgefunden hat (beziehe dich dabei explizit auf die bearbeiteten Quellen bzw. deine Notizen dazu).3. Bewerte abschliessend, inwiefern (aus heutiger Sicht) die Haltungen der sich selbst als deutsch und polnisch bezeichnenden Gruppierungen nachvollziehbar sind. Beziehe ggf. auch aktuelle Beispiele mit ein.

Tab. 2: Materialgrundlage und Aufgabenstellungen

Die Auswertung der auf Grundlage dieser Aufgabenstellungen entstandenen Texte erfolgte im Rahmen einer evaluativen qualitativen Inhaltsanalyse nach Kuckartz (2018, 123–140). Im Sinne einer Operationalisierung als Messbarmachung eines theoretischen Konzeptes (Döring & Bortz, 2016, 228) sind dazu zunächst Indikatoren zu benennen, anhand derer historische Denkleistungen in Texten identifiziert und ausgewertet werden können. Ausgehend von der Beschreibung der «mentalen Operationen der Sinnbildung» (Rüsen, 2013, 35) als Wahrnehmung, Deutung, Orientierung und Motivation⁸ werden auf struktureller Ebene die fünf Bereiche der Quellenbezüge, Bezüge zu externen Wissensbeständen, vergangenheitsbezogenen Deutungsleistungen sowie der gegenwarts- bzw. zukunftsbezogenen Deutungsleistungen und der Identifikation historischer Aktanten benannt, mit denen der Anspruch einhergeht, historische Denkleistungen auf der Satzebene als Kodiereinheit (Kuckartz, 2018, 41–44) erschöpfend zu erfassen, wobei Doppelcodierungen möglich und üblich sind. Sätze mit Quellenbezügen sind dabei jene, in denen Schreibende explizit oder implizit auf die zur Verfügung gestellten Quellen rekurrieren. In Abgrenzung davon werden unter Bezügen zu externen Wissensbeständen sämtliche Verweise auf andere unterrichtliche oder ausserunterrichtliche Informationsquellen verstanden. Deutungen bezeichnen hier Sätze, die dadurch einen narrativierenden Charakter haben, dass sie Ereignisse in der Vergangenheit sinnbildend miteinander verbinden (vergangenheitsbezogene Deutungsleistungen) bzw. vergangene Ereignisse mit der Gegenwart, z. B. mit gegenwärtigen Wertmassstäben und Positionen, verknüpfen (gegenwarts- bzw. zukunftsbezogene Deutungsleis-

8 Der Aspekt der Motivation wird hier als aussernarratives Handeln als Konsequenz des historischen Denkprozesses (Rüsen, 2013, 36) nicht weiter verfolgt.

tungen).⁹ Bei der Identifikation historischer Aktanten werden abschliessend auf Wortebene als Kodiereinheit die Träger und Ziele historischer Prozesse markiert, die von den Schreibenden benannt werden.

Die Analyse der Qualität der identifizierten Manifestationen historischer Denkleistungen erfolgt auf Grundlage eines auf den Triftigkeiten nach Rüsen basierenden Analyserasters (Rüsen, 1990, 87–91).¹⁰ Diese Kriterien, anhand derer Erzählungen im oben dargelegten Sinne hinsichtlich ihrer Zustimmungsfähigkeit und damit auch ihres narrativen Wahrheitsanspruchs überprüft werden können (Rüsen, 2013, 57–58), sind in der Geschichtsdidaktik bereits ausgiebig diskutiert worden.¹¹ Relativer Konsens herrscht hierbei bei der Definition der empirischen Triftigkeit, die sich nach Rüsen in einem «überprüfbareren Tatsachenbezug» (ebd., 60) niederschlägt, der durch die Offenlegung einerseits der der Narration zugrundeliegenden Quellen sowie ihrer Provenienz, andererseits der Auswahlentscheidungen der Quellenkritik, die zur Verwendung genau dieser Quellen geführt haben, gewährleistet wird. Neben die Auseinandersetzung mit der Perspektivgebundenheit der Quellen tritt die Reflexion sowohl des Standpunktes, von dem aus erzählt wird, als auch der leitenden «Hinsichten» (Rüsen, 1986, 104) im Sinne mehr oder weniger objektiver Beurteilungsmassstäbe, die Rüsen unter dem Aspekt der normativen Triftigkeit zusammenfasst (Rüsen, 2013, 61–62). Erst durch deren Verknüpfung mit der Gegenwart kann der Erzählung ein Sinn im Sinne einer Orientierungsfunktion für gegenwärtige Menschen zukommen (Bergmann, 2008, 144). Ebendiese bildet den Kern der bei Rüsen eher vage bleibenden Kategorie der narrativen Triftigkeit, in deren Rahmen die Gültigkeitsbedingungen von Narrationen sowie ihre Eignung für Perspektiv-erweiterungen zu untersuchen sind (Rüsen, 1983, 114).

Ausgehend von den vorgestellten Definitionen von Triftigkeit werden die so identifizierten strukturellen Merkmale historischer Denkleistungen hinsichtlich ihrer Ausprägungen evaluiert. Mit Blick auf die Treatmentgruppen besteht

9 Zum Begriff der Deutung vgl. mit direktem Bezug zu historischem Lernen Rüsen (1989, 99). Vom Begriff des Urteils wird hier Abstand genommen, da unter Deutungen auch Aussagen verstanden werden, die die Kriterien des Urteils z. B. nach Becker (2017, 322–323) (noch) nicht erfüllen.

10 Rüsen (2013, 57–62) spricht nicht mehr von Triftigkeiten, sondern von Plausibilitäten. Hier wird dennoch weiterhin der synonyme und etablierte Begriff der Triftigkeit verwendet, die synonymen Begriffsbedeutungen beschreibt z. B. Körber (2018, 22).

11 Exemplarisch für Operationalisierungen der Triftigkeitskriterien sei hier nur verwiesen auf die Arbeiten von Waldis, Marti und Nitsche (2015), Sochatzy, Schöner und Schreiber (2013) sowie Hasberg (2007).

die Grundannahme der Auswertung darin, dass die verwendeten Medien «[m]it ihren medienspezifischen Eigenschaften [...] diese Sinnbildungsprozesse» (Hodel, 2016, 155) prägten. Die Lernertexte sind dementsprechend als Resultat der Kombination individueller historischer Denkleistungen und Medieneigenschaften zu betrachten (ebd., 157). Dem Rechercheprozess der Lernenden in den Digitalarchiven kommt dabei eine Sonderrolle zu, da es sich dabei ebenfalls um eine Kombination wie die gerade beschriebene handelt, bei der die Lernenden sowohl im Rahmen der Entscheidung für adäquate Suchbegriffe und der Auswahl zielführender Suchparameter als auch bei der Selektion der Suchresultate zumindest implizit auf Triftigkeitskriterien im beschriebenen Sinne zurückgreifen müssen (ebd., 157–158). Damit führt die Recherche die Teilbereiche des historischen Denkens und der (historischen) Medienkompetenz zusammen,¹² wobei davon ausgegangen wird, dass sich dieser ebenfalls (auch) triftigkeitsbasierte Prozess in der konkreten Ausgestaltung der Texte, die auf seiner Grundlage entstehen, so niederschlägt, dass er sich wiederum triftigkeitsanalytisch nachweisen lässt. Erwartet wird, dass u. a. der Einblick in die Vielzahl potenzieller Quellen zu einer Sensibilisierung hinsichtlich divergierender leitender Hin-sichten im Bereich der normativen Triftigkeit sowie einer vertieften Reflexion der Gültigkeit von Aussagen zu historischen Sachverhalten im Sinne der narrativen Triftigkeit beiträgt.

Die Triftigkeiten sind in Subcodes operationalisiert worden, die auf einer abgestuften Ratingskala von eins bis drei vergeben werden. Dies sei hier exemplarisch anhand des Merkmals «vergangenheitsbezogene Deutungsleistungen» entfaltet.

12 Dementsprechend beschreibt Kerber (2017, 73) in seiner Modellierung einer historischen Medienkompetenz die Recherche als Querschnittsaufgabe, die sich aus deren geschichtsdidaktischen Inhalts- sowie «Handlungsfelder[n] geschichtsunterrichtlicher Medienpädagogik» (Kerber, 2017, 53) speist.

Denkleistung	Vergangenheitsbezogene Deutungsleistung (VD)		
Triftigkeit	Prüfung der empirischen Triftigkeit	Prüfung der normativen Triftigkeit	Prüfung der narrativen Triftigkeit
Subcode (Kürzel)	Qualität vergangenheitsbezogener Deutungsleistungen (VDRelGM)	Qualität leitender Hinsichten (VLHRelGM)	Reichweite vergangenheitsbezogener Deutungsleistungen (VRRelGM)
Leitfrage	Wie intersubjektiv nachvollziehbar wird die vergangenheitsbezogene Deutungsleistung ausgeführt?	Auf welcher Grundlage werden Deutungen historischer Ereignisse vorgenommen?	Mit welchem Gültigkeitsanspruch werden vergangenheitsbezogene Deutungsleistungen formuliert?
Ratings	intransparent (1) – transparent (2) – transparent und belegt (3)	intransparent (1) – transparent (2) – wertend (3)	intransparent (1) – transparent (2) – wertend (3)
Ankerbeispiel (Dokumentchiffre Erhebungszeitpunkt) (Rating)	Theoretisch waren die Bedingungen für eine Integration gegeben, das Problem waren nur die preussischen Einwohner und die Regierung, da diese die Polen eindeutschen wollten, so dass diese ihre Kultur aufgeben. (AnWorg24 1) (2)	Eine solche Integration wurde im Zusammenhang mit den polnischen Migranten nicht angestrebt. (YvMaen20 2) (2)	In ihre Zeitung «Wiarus Polski», welche auf polnisch verfasst ist, versuchen sie heranzukommen und sie an ihre tolle Kultur und an sie selbst zu erinnern. (SaErum04 1) (1)

Tab. 3: Subcodes und Ratings zum Code «vergangenheitsbezogene Deutungsleistungen»

4 Auswertung am Beispiel des Teilbereichs «vergangenheitsbezogene Deutungsleistungen»

Die Codiererergebnisse der qualitativen Inhaltsanalyse wurden einer Inter- und Intracoderüberprüfung unterzogen. Für erstere haben zwei an der Studie beteiligte Hilfskräfte, die durch den Verfasser geschult worden sind, zehn Prozent der Gesamtstichprobe eigenständig codiert. Für die Identifikation der Manifestationen historischer Denkleistungen konnte ein κ nach Brennan und Prediger (1981, 692–694) von .90 bis .96 ($M = .92$) erreicht werden; über das gesamte Codesystem ergab sich mit $M\kappa = .79$ ein zufriedenstellende Wert (Kuckartz, 2018, 210). Für die

Intracoderüberprüfung wurden nach Abschluss des Codierprozesses acht Dokumentenpaare und damit ca. 14,5 % der Gesamtstichprobe zufällig ausgewählt und erneut codiert. Hier ergaben sich mit Werten zwischen $\kappa = .95$ und $\kappa = .98$ ($M = .97$) für die Identifikation der Manifestationen historischer Denkleistungen sowie von $M\kappa = .88$ für das gesamte Codesystem ebenfalls befriedigende Ergebnisse.

Für die weitere Auswertung sind die Resultate der qualitativen Inhaltsanalyse quantifiziert worden, um sie einer inferenzstatistischen Untersuchung zugänglich zu machen und auf diesem Wege relevante Unterschiede zwischen den Erhebungszeitpunkten und Treatmentgruppen zu identifizieren. Die absolute Anzahl der Manifestationen historischer Denkleistungen pro Text¹³ N_x wurde zur Berechnung relativer Codehäufigkeiten in Bezug zu den Kennwerten Satzzahl bzw. Wörterzahl gesetzt:

$$X_{Rel} = \frac{N_x}{N_{Satz/Wörter}}; 0 \leq X_{Rel} \leq 1$$

Zur Bestimmung der Qualität der identifizierten Elemente historischer Denkleistungen sind die relativen Häufigkeiten der einzelnen Ratings zu einem gewichteten Mittel¹⁴ zusammengefasst worden, bei dem die Ratingstufen (1–3) als Gewichte fungierten.

$$X_{RelGM} = \frac{(X_{RelR1} * 1) + (X_{RelR2} * 2) + (X_{RelR3} * 3)}{6}$$

Mit Blick auf die erste Forschungsfrage ist die Leistungsveränderung im Hinblick auf die relative Häufigkeit und die Umsetzung vergangenheitsbezogener Deutungsleistungen innerhalb der zu untersuchenden Gruppen, also denjenigen Lernenden, die mit didaktisierten Quellen gearbeitet haben (Q , $n = 14$), sowie der Gesamtheit all jener, die mit den verschiedenen Formen historischer Zeitungen gearbeitet haben (Z , $n = 41$), zu prüfen. Unter anderem Dimitrov und Rumrill

13 Im Hinblick auf die Auswertung der Absolutwerte ist die Frage zu diskutieren, inwiefern die bloße, möglicherweise redundante Anzahl der Manifestationen historischer Denkleistungen in längeren Texten einen relevanten Qualitätsunterschied im Vergleich zu kürzeren Texten mit damit auch zwangsläufig weniger entsprechenden Manifestationen darstellt. Vor diesem Hintergrund wird hier von einer Darstellung der Absolutwerte Abstand genommen, im Kontext der Gesamtstudie werden diese jedoch ausführlich berücksichtigt.

14 Die mathematischen Eigenschaften des gewichteten Mittels stellt z. B. Bullen (2003, 63) dar.

(2003, 160) schlagen insgesamt vier Vorgehensweisen für den Vergleich von Gruppen in Prätest-Posttest-Designs vor, wobei sie insbesondere die Analyse von Differenzscores sowie die Durchführung von Kovarianzanalysen (ANCOVA) als relevante Methoden herausstellen (ebenso auch bei van Breukelen, 2006, 920). Auch wenn die grossen Vorbehalte, die gerade in älterer Literatur gegen die Verwendung von Differenzscores im Hinblick auf ihre mangelhafte Reliabilität angeführt werden (Cronbach & Furby, 1970, 78), in aktuelleren Diskussionen so nicht mehr vertreten werden (z. B. Kim & Steiner, 2021, 1361–1365), wird hier insbesondere mit Blick auf die zugrundeliegenden Vorannahmen¹⁵ der ANCOVA als Messinstrument der Vorzug gerade für nicht randomisierte Studien gegeben (z. B. Dimitrov & Rumrill, 2003, 163). Daher wird für alle Vergleiche, bei denen die Voraussetzungen für eine ANCOVA (Bortz & Schuster, 2010, 311–314) erfüllt sind, dieses Verfahren ausgewählt, andernfalls wird ein Gruppenvergleichstest über die Differenzscores $t_2 - t_1$ der Variablen berechnet. Im Folgenden werden die Ergebnisse der Berechnung von ANCOVAs über die Posttestscores der Codes zu vergangenheitsbezogenen Deutungsleistungen (Innersubjektfaktor: Gruppenzugehörigkeit; Kovariate: Prätestscore) berichtet. Die unadjustierten deskriptiven Werte, die um das Prätestergebnis bereinigten Werte sowie die Resultate der ANCOVAs sind in Tabelle 3 zusammengefasst:

15 Die Vorannahmen der ANCOVA, insbesondere die Regression zur Mitte, sowie die Untersuchung von Differenzwerten, die Annahme gleicher Veränderungen bei gleichem Treatment, beschreibt z. B. van Breukelen (2006, 924–925).

Variable vergangenheitsbezogene Deutungsleistungen (Rel.) VDRel $F(1, 52) = 5.30, p = .025, \text{partielles } \eta^2 = .09$					
unadjustiert				adjustiert	
Gruppe	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Q	14	.39	.15	.41	.04
Z	41	.53	.17	.53	.03

Variable Qualität vergangenheitsbezogener Deutungsleistungen (Rel.) VDRelgM gewichtetes Mittel $F(1, 52) = 5.17, p = .027, \text{partielles } \eta^2 = .09$					
unadjustiert				adjustiert	
Gruppe	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Q	14	.13	.05	.13	.02
Z	41	.18	.06	.17	.01

Variable Qualität leitender Hintersichten (Rel.) VLHRelgM gewichtetes Mittel $F(1, 52) = 5.16, p = .025, \text{partielles } \eta^2 = .09$					
unadjustiert				adjustiert	
Gruppe	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Q	14	.12	.05	.13	.02
Z	41	.17	.07	.17	.01

Variable Reichweite vergangenheitsbezogener Deutungsleistungen (Rel.) VRRelgM gewichtetes Mittel $F(1, 52) = 4.38, p = .041, \text{partielles } \eta^2 = .08$					
unadjustiert				adjustiert	
Gruppe	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Q	14	.08	.03	.09	.01
Z	41	.12	.06	.12	.01

Tab. 4: Unadjustierte und adjustierte deskriptive Statistiken sowie Ergebnisse der ANCOVAs zum gruppenweisen (Q, Z) Vergleich der Posttestscores (Forschungsfrage 1)

Die Ergebnisse der ANCOVAs zeigen durchweg einen signifikanten Unterschied¹⁶ mit mittlerem Effekt nach Cohen (2013, 79–81) zwischen den Posttestergebnissen der Z- und jenen der Q-Gruppe, wobei die Ergebnisse der Zeitungsgruppen sowohl im Hinblick auf die relative Häufigkeit vergangenheitsbezogener

¹⁶ Als Signifikanzniveau bei den durchgeführten zweiseitigen Tests wird hier und im Folgenden $p < 0,05$ verwendet. Es sei betont, dass aufgrund der geringen Stichprobengröße und des Designs Signifikanz als Hinweis auf eine statistische Tendenz verstanden wird, die unter Vorbehalt zu interpretieren ist.

Deutungsleistungen als auch auf die definierten Triftigkeitsmerkmale derselben im Mittel durchweg höher ausgefallen sind. Dementsprechend scheint es berechtigt, einen Zusammenhang zwischen diesen Resultaten und dem Einsatz der verschiedenen Formen historischer Zeitungen anzunehmen.

Zur konkreten Interpretation des Befunds ist es mit Blick auf die zweite Forschungsfrage notwendig, den Leistungszuwachs der einzelnen Gruppen, die mit historischen Zeitungen gearbeitet haben (Recherche in Digitalarchiven DA, $n = 17$; Faksimiles historischer Zeitungen in digitalisierter Form ZD, $n = 12$; gedruckte Faksimiles historischer Zeitungen $n = 12$, ZA) zu vergleichen. Auch dazu wurden ANCOVAs über die Posttestscores der beschriebenen Codes (Innersubjektfaktor: Zugehörigkeit Zeitungsgruppe; Kovariate: Prätestscore) berechnet. Die Ergebnisse sind in Tabelle 4 aufgeführt. Da die Scores der Variable zur Reichweite vergangenheitsbezogener Deutungsleistungen die Voraussetzungen der ANCOVA nicht erfüllten,¹⁷ wurde für diese Variable ein Kruskal-Wallis-H-Test (Kruskal & Wallis, 1952, dazu auch Bortz & Lienert, 2008, 222–228) über die Differenzscores $t_2 - t_1$ berechnet,¹⁸ der in Tabelle 5 berichtet wird.

17 Die Regressionssteigungen der Stichprobenscores sind für diese Variable nicht gegeben, darüber hinaus sind auch die Residuen nicht normalverteilt, was zusammen bei unterschiedlich grossen Stichproben nach Bortz und Schuster (2010, 311) dazu führt, dass die ANCOVA kontraindiziert ist.

18 Von der Verwendung einer ANOVA wurde aufgrund einiger Ausreisser in zwei der drei Gruppen Abstand genommen.

Variable vergangenheitsbezogene Deutungsleistungen (Rel.) VDRel $F(2, 37) = 5.47, p = .008, \text{partielles } \eta^2 = .23$					
Gruppe	unadjustiert			adjustiert	
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
DA	17	.45	.15	.45	.03
ZD	12	.63	.14	.64	.03
ZA	12	.54	.18	.54	.04

Variable Qualität vergangenheitsbezogener Deutungsleistungen (Rel.) VDRelgM gewichtetes Mittel $F(2, 37) = 57.35, p = .002, \text{partielles } \eta^2 = .28$					
Gruppe	unadjustiert			adjustiert	
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
DA	17	.16	.05	.14	.01
ZD	12	.21	.07	.23	.02
ZA	12	.17	.17	.17	.02

Variable Qualität leitender Hinsichten (Rel.) VLHRelgM gewichtetes Mittel $F(2, 37) = 3.86, p = .030, \text{partielles } \eta^2 = .17$					
Gruppe	unadjustiert			adjustiert	
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
DA	17	.15	.05	.14	.01
ZD	12	.19	.06	.20	.01
ZA	12	.17	.08	.19	.01

Tab. 5: Unadjustierte und adjustierte deskriptive Statistiken sowie Ergebnisse der ANCOVAS zum gruppenweisen (DA, ZD, ZA) Vergleich der Posttestscores (Forschungsfrage 2)

Variable Reichweite vergangenheitsbezogener Deutungsleistungen (Rel.) VRRelgM.diff Differenzscore ($t_2 - t_1$) der gewichteten Mittel $\chi^2(2) = 8.88, p = .01219$			
Gruppe	<i>n</i>	<i>M</i>	<i>SD</i>
DA	17	.00	.05
ZD	12	.06	.05
ZA	12	-.01	.07

Tab. 6: Deskriptive Statistiken sowie Ergebnisse des Kruskal-Wallis-H-Tests zum gruppenweisen (DA, ZD, ZA) Vergleich der Differenzscores (Forschungsfrage 2)

19 Die Verwendung der χ^2 -Verteilung an dieser Stelle sowie den Verzicht auf die Angabe einer Effektstärke für den Omnibustest erläutern Rasch, Friese, Hofmann und Naumann (2021, 120–121).

Die Übersicht zeigt, dass sich die Posttestscores der Gruppen, die mit historischen Zeitungen gearbeitet haben, untereinander signifikant mit starkem Effekt voneinander unterscheiden. Die für die Ergebnisse der ANCOVAS Bonferroni-korrigierte Post-hoc-Analyse ergab einen signifikanten Unterschied zwischen der DA- sowie der ZD-Gruppe für alle drei getesteten Variablen (VDRel: $p = .006$, $M_{\text{Diff}} = -.19$; VDRelGM: $p = .002$, $M_{\text{Diff}} = -.08$; VLHGGM: $p = .043$, $M_{\text{Diff}} = -.06$). Für die Variable zur Qualität vergangenheitsbezogener Deutungsleistungen wurde zudem der Unterschied zwischen den Gruppen DA und ZA signifikant ($p = .045$, $M_{\text{Diff}} = .06$).

Für die Post-hoc-Analyse des Kruskal-Wallis-H-Tests über die Variable zur Reichweite vergangenheitsbezogener Deutungsleistungen wurden Dunn-Bonferroni-Tests (Dunn, 1964) berechnet. Diese ergaben signifikante Unterschiede zwischen den Gruppen DA und ZD ($z = -2.72$, $p = .020$, $r = .51$) sowie zwischen ZA und ZD ($z = 2.50$, $p = .038$, $r = .51$)

5 Diskussion und Fazit

Zusammengefasst ergibt sich damit für diesen Ausschnitt aus der Gesamtheit der untersuchten Variablen ein recht deutliches Bild. Dabei ist jedoch zu beachten, dass an dieser Stelle die Wechselwirkungen zwischen den verschiedenen Manifestationen historischer Denkleistungen sowie zwischen ihren qualitativen Ausprägungen nicht mitberücksichtigt werden können, obwohl ihnen für eine umfassende Interpretation der vorgestellten Ergebnisse zum Teilbereich der vergangenheitsbezogenen Deutungsleistung eine nicht zu unterschätzende Bedeutung zukäme. Darüber hinaus sind im Rahmen der Fragebogen- und Beobachtungsstudie erhobene potenzielle Störeinflüsse auf die gewonnenen Ergebnisse, wie beispielsweise der Einfluss der verschiedenen Lehrkräfte oder das grundsätzliche Interesse der Lernenden an Geschichte, noch nicht in die konkrete Auswertung aufgenommen worden. Als zentrale Ergebnisse sind vor diesem Hintergrund die beiden folgenden Aspekte festzuhalten.

Erstens hat der Gruppenvergleich zwischen den Zeitungsgruppen und der Q-Gruppe (Forschungsfrage 1) gezeigt, dass erstere relativ zur Textlänge sowohl im Mittel signifikant mehr entsprechende Deutungsleistungen formuliert als auch diese besser belegt bzw. sicherer argumentativ in den Gesamttext eingebunden hat (Variable VDRelGM, Merkmal der empirischen Triftigkeit). Darüber hinaus konnten die Schreibenden der Z-Gruppen die leitenden Hinsichten, auf deren Grundlage sie ihre Deutungsleistungen formulierten, deutlicher offen-

legen (Variable VLHRelgM, Merkmal der normativen Triftigkeit) sowie deren Gültigkeitsansprüche in Grundzügen in ihren Texten reflektieren (Variable VRRRelgM, Merkmal der narrativen Triftigkeit). Der damit angedeutete Zusammenhang zwischen dem Einsatz historischer Zeitungen und dem Umgang mit vergangenheitsbezogenen Deutungsleistungen wurde in den Untersuchungen zu Forschungsfrage 2 konkretisiert.

Die entsprechenden Vergleiche für die Digitalarchive-, Digitalisate- und Faksimiledrucke-Gruppen haben nachgewiesen, dass es zu einem signifikanten Unterschied mit durchweg starkem Effekt auch zwischen den Zeitungsgruppen gibt. Zum anderen konnte durch die Post-hoc-Analysen gezeigt werden, dass die signifikanten Unterschiede für alle untersuchten Teilbereiche zwischen derjenigen Gruppe, die mit digitalisierten Zeitungsfaksimiles gearbeitet, und derjenigen, die eigenständig in Digitalarchiven nach Zeitungen recherchiert hat, bestehen. Für die Untersuchungen der Merkmale empirischer und narrativer Triftigkeit ergaben sich derartige Unterschiede zudem auch zwischen ersterer und der Gruppe, in der Faksimiledrucke historischer Zeitungen zum Einsatz gekommen sind, wobei mit Blick auf die deskriptiven Werte diejenigen Lernenden, die Zeitungsdigitalisate verwendet haben, durchweg die nach den Massstäben der Auswertung besseren Ergebnisse produziert haben.

Im Rahmen der untersuchten Stichprobe zeichnet sich damit als Tendenz ab, dass gerade der Einsatz von digitalisierten Zeitungsfaksimiles in einem positiven Zusammenhang mit der Art und Weise steht, auf die Lernende vergangenheitsbezogene Deutungsleistungen gestalten. Im Gegensatz dazu scheinen die insbesondere auf den Bereich der narrativen Triftigkeit ausgerichteten Ziele, die auf theoretischer Ebene für die Recherche in digitalen Zeitungsportalen angenommen worden sind, in keiner Weise erreicht worden zu sein. Im Gegenteil liegen die deskriptiven Werte hier durchweg unter denjenigen der anderen Zeitungsgruppen bzw. für den Bereich der Reichweite der vergangenheitsbezogenen Deutungsleistungen auf nahezu dem gleichen Niveau wie die Faksimiledrucke-Gruppe. Diese Konstellation ist abschliessend insbesondere mit Blick auf die Verbindung des Mediums Zeitung mit der Medialität des Digitalen zu diskutieren.

Die Gruppenvergleiche im Rahmen der ersten Forschungsfrage haben als Tendenz aufgezeigt, dass der Einsatz historischer Zeitungen in den verschiedenen Darreichungsformen die bislang theoretisch formulierten Ansprüche zumindest im Rahmen dieser Untersuchung in Bezug auf die Formulierung vergangenheitsbezogener Deutungsleistungen erfüllen zu können scheint. Auch im

Einzelvergleich der deskriptiven Kennwerte übertreffen die Ergebnisse der Zeitungsgruppen in diesem Bereich diejenigen der Q-Gruppe. Die inferenzstatistische Untersuchung, inwiefern diese Unterschiede der Posttestscores signifikant sind, wird an anderer Stelle zu berichten sein. Zu fragen bleibt aber, warum die relativ simple Form des Digitalisats der historischen Zeitung eine relativ deutliche positive Tendenz aufweist, während die Recherchetätigkeit im digitalen Raum tendenziell zu im Durchschnitt der Zeitungsgruppen eher schlechteren Resultaten führt. Als zentraler Faktor soll hier das Verhältnis von Komplexität und Nützlichkeit der jeweiligen medialen Form der historischen Zeitung zur Erklärung herangezogen werden. So standen einerseits die Lernenden, die mit gedruckten Faksimiles zu arbeiten hatten, vor der Herausforderung, in grossen Material-, d. h. hier Papiermengen, einen Überblick zu gewinnen und zu behalten. Die Möglichkeiten der digitalen Endgeräte (z. B. Lesezeichen setzen, Screenshots erstellen usw.) standen ihnen nicht zur Verfügung. Andererseits muss in Betracht gezogen werden, dass die Recherchetätigkeit in den Digitalarchiven einen so grossen Teil der Aufmerksamkeit der Lernenden in Anspruch genommen hat, dass es ihnen nicht mehr möglich war, mit dem gefundenen Material ähnlich der ZD-Gruppe zu arbeiten oder gar den Prozess der Recherche und den Einblick in den Fundus potenziell zu verwendender Quellen in das Schreibergebnis einfließen zu lassen. An dieser Stelle ist auch das verkürzte Treatment zu berücksichtigen, das möglicherweise nicht ausgereicht hat, um die Lernenden in einer Weise mit den entsprechenden digitalen Heuristiken vertraut zu machen, um die theoretisch erwarteten Ergebnisse zu erreichen.

Es sei damit hier als abschliessendes Fazit und zugleich als These formuliert, dass die identifizierten Unterschiede einerseits, wie beschrieben, mit den geschichtsdidaktischen Potenzialen der historischen Zeitung in Verbindung stehen. Andererseits scheint die Vereinfachung des Arbeitsprozesses, die mit der Verwendung der Zeitungsdigitalisate einhergeht, die Verfügbarkeit dieser Potenziale für die Lernenden zu erhöhen. Im Gegensatz dazu liegt die Annahme nahe, dass die Komplexitätssteigerung, die aus den Rechercheaufgaben resultierte, im durchgeführten Setting zu einer partiellen Überforderung beigetragen hat, die der Wirksamkeit des Mediums Zeitung für das historische Lernen tendenziell abträglich gewesen ist. In der weiteren Auswertung der erhobenen Daten ist nun zu überprüfen, inwiefern sich derartige Beobachtungen auch im Hinblick auf die übrigen Manifestationen historischer Denkleistungen machen lassen. Darüber hinaus dürfte von höchstem Interesse sein, im Zuge der Auswertung der die Haupterhebung flankierenden Fragebogenstudie zu untersu-

chen, ob die Einschätzung der Verwendung des Mediums Zeitung seitens der Lernenden einerseits und der medialen Darreichung andererseits die aufgestellte These der Komplexitätsreduktion und Nützlichkeit durch den Einsatz der Zeitungsdigitalisate stützt. Festgehalten werden darf aber bis hierhin und unter Berücksichtigung aller dargelegten Einschränkungen der Aussagekraft der Studie, dass die vorgestellten Ergebnisse als Hinweis darauf gelesen werden können, dass der historischen Zeitung ein tatsächliches Potenzial für historisches Lernen zukommen kann.

Literatur

- Ankersmit, F. (1994). *History and tropology: The Rise and fall of metaphor* ([ACLS Humanities E-Book edition]). Berkeley: University of California Press. Abgerufen von <http://ark.cdlib.org/ark:/13030/kt9k4016d3/>.
- Becker, Axel (2017). Historische Urteilsbildung. In Michele Barricelli & Martin Lücke (Hrsg.), *Handbuch Praxis des Geschichtsunterrichts* (2. Aufl., Bd. 1, S. 316–325). Schwalbach/Ts.: Wochenschau.
- Bergmann, Klaus (2008). *Geschichtsdidaktik: Beiträge zu einer Theorie historischen Lernens* (3. Aufl.). Schwalbach/Ts.: Wochenschau.
- Bernsen, Daniel (2017). Online recherchieren. In Daniel Bernsen & Ulf Kerber (Hrsg.), *Praxishandbuch Historisches Lernen und Medienbildung im digitalen Zeitalter* (S. 313–319). Bonn: BPB.
- Bortz, Jürgen / Lienert, Gustav A. (2008). *Verteilungsfreie Methoden in der Biostatistik* (3. Aufl.). Heidelberg: Springer.
- Bortz, Jürgen / Schuster, Christof (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Berlin: Springer.
- Bösch, Frank (2019). Zeitungs-Geschichte: Zum Wandel eines Mediums. *Praxis Geschichte*. (2), 4–9.
- Brennan, Robert L. & Prediger, Dale J. (1981). Coefficient Kappa: Some uses, missuses, and alternatives. *Educational and Psychological Measurement*, 41(3), 687–699.
- Broekelen, Gerard J. P. van (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, 59(9), 920–925.
- Bullen, P. S. (2003). *Handbook of means and their inequalities*. Boston: Kluwer Academic.
- Cohen, Jacob (2013). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hoboken: Taylor and Francis.

- Cronbach, Lee J. & Furby, Lita (1970). How we should measure «change» – or should we? *Psychological Bulletin*, 74(1), 68–80.
- Dimitrov, Dimitar M. & Rumrill, Phillip D. (2003). Pretest-posttest designs and measurement of change. *Work*, 20(2), 159–165.
- Döring, Nicola & Bortz, Jürgen (2016), *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Berlin: Springer.
- Droysen, Johann G. (1868). *Grundzüge der Historik*. Leipzig: Velt.
- Dunn, Olive J. (1964). Multiple comparisons using rank sums. *Technometrics*, 6(3), 241. <https://doi.org/10.2307/1266041>
- Fleischhacker, Michael (2014). *Die Zeitung: Ein Nachruf*. Wien: Brandstätter.
- Föhr, Pascal (2017). *Historische Quellenkritik im Digitalen Zeitalter*. Abgerufen von https://edoc.unibas.ch/64111/1/F%C3%B6hr_Pascal-Historische_Quellenkritik_im_Digitalen_Zeitalter-2018.pdf.
- Gau «Ruhr und Lippe» des Alldeutschen Verbandes (Hrsg.) (1901). *Die Polen im rheinisch-westfälischen Steinkohlen-Bezirk: mit einem statistischen Anhang, einer Sammlung polnischer Lieder und zwei Karten*. München: Lehmann.
- Goldman, Susan R. (2004). Cognitive aspects of constructing meaning through and across multiple texts. In Nora Shuart-Faris & David Bloome (Hrsg.), *Uses of intertextuality in classroom and educational research* (S. 313–347). Greenwich: Information Age Pub.
- Groth, Otto (1960). *Das Wesen des Werkes*. Berlin: de Gruyter.
- Hagenauer, Gerade (2010). Kurzzeitinterventionen versus Langzeitinterventionen. In Tina Hascher & Bernhard Schmitz (Hrsg.), *Pädagogische Interventionsforschung: Theoretische Grundlagen und empirisches Handlungswissen* (S. 243–251). Weinheim, München: Juventa.
- Hasberg, Wolfgang (2007). Die Entzauberung der Hrosvith von Gandersheim – oder: Dekonstruktion als Akt entdeckenden historischen Lernens. In Wolfgang Hasberg & Wolfgang E. J. Weber (Hrsg.), *Geschichte entdecken: Karl Filser zum 70. Geburtstag* (S. 211–242). Berlin: LIT.
- Henning, Hansjoachim, & Sieg, Uwe (1997). Quellensammlung zur Geschichte der deutschen Sozialpolitik 1867 bis 1914: IV. Abteilung (1905–1914) (Bd. 3). Darmstadt: Wissenschaftliche Buchgesellschaft. Abgerufen von <http://www.urn.fi/urn:nbn:de:he-bis:34-2018040555125>.
- Hodel, Jan (2016). Geschichte – Medien – Suche: Kompetenter Umgang mit komplexen Herausforderungen. In Saskia Handro & Bernd Schönemann (Hrsg.), *Aus der Geschichte lernen?: Weisse Flecken der Kompetenzdebatte* (S. 153–168). Berlin: LIT.
- Kerber, Ulf (2017). Historische Medienbildung – ein transdisziplinäres Modell für den Geschichtsunterricht. In Daniel Bernsen & Ulf Kerber (Hrsg.), *Praxishandbuch Historisches Lernen und Medienbildung im digitalen Zeitalter* (S. 45–82). Bonn: BPB.

- Kim, Yongnam & Steiner, Peter M. (2021). Gain scores revisited: A graphical models perspective. *Sociological Methods & Research*, 50(3), 1353–1375. <https://doi.org/10.1177/0049124119826155>
- Körper, Andreas (2018). Geschichte – Spielen – Denken: Kontingenzverschiebungen und zweiseitige Triftigkeiten. *Medienpädagogik Zeitschrift für Theorie und Praxis der Medienbildung*, 19–35.
- Krameritsch, Jakob (2007). *Geschichte(n) im Netzwerk: Hypertext und dessen Potenziale für die Produktion, Repräsentation und Rezeption der historischen Erzählung*. Münster: Waxmann.
- Kruskal, William H. & Wallis, W.A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583. <https://doi.org/10.2307/2280779>
- Kuchler, Christian (2016). Zwischen geschichtsdidaktischer Forschung und aktuellem Unterrichtseinsatz: Historisches Lernen und Zeitung. In Christian Kuchler & Benjamin Städter (Hrsg.), *Zeitungen von gestern für das Lernen von morgen? Historische Tagespresse im Geschichtsunterricht* (S. 31–49). Göttingen.
- Kuckartz, Udo (2018). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung* (4. Aufl.). Weinheim: Beltz Juventa.
- Kultusministerkonferenz (2005). Einheitliche Prüfungsanforderungen in der Abiturprüfung: Geschichte. Abgerufen von https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Geschichte.pdf am 08. 11. 2022.
- Leutner, Detlev (2012). Perspektiven pädagogischer Interventionsforschung. In Eckart Severing & Reinhold Weiss (Hrsg.), *Qualitätsentwicklung in der Berufsbildungsforschung* (1. Aufl., S. 17–28). Bielefeld: Bertelsmann.
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (2014). Kernlehrplan für die Sekundarstufe II Gymnasium / Gesamtschule in Nordrhein-Westfalen: Geschichte. Abgerufen von https://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_SII/ge/KLP_GOSt_Geschichte.pdf am 08.11.2022.
- Rasch, Björn / Friese, Malte / Hofmann, Wilhelm & Naumann, Ewald (2021). *Quantitative Methoden 2: Einführung in die Statistik für Psychologie, Sozial- & Erziehungswissenschaften* (5. Aufl. 2021). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Rüsen, Jörn (1983). *Historische Vernunft: Die Grundlagen der Geschichtswissenschaft*. Göttingen: V & R.
- Rüsen, Jörn (1986). *Rekonstruktion der Vergangenheit: Die Prinzipien der historischen Forschung*. Göttingen: V & R.
- Rüsen, Jörn (1989). *Lebendige Geschichte: Foren und Funktionen des historischen Wissens*. Göttingen: V & R.

- Rüsen, Jörn (1990). *Zeit und Sinn: Strategien historischen Denkens*. Frankfurt am Main: Fischer.
- Rüsen, Jörn (2013). *Historik: Theorie der Geschichtswissenschaft*. Köln: Böhlau.
- Sauer, Michael (2006). «Was sich begeben und zugetragen hat»: Die Zeitung als Quelle im Geschichtsunterricht. In Markus Bernhardt, Gerhard Henke-Bockschatz, & Michael Sauer (Hrsg.), *Bilder – Wahrnehmungen – Konstruktionen: Reflexionen über Geschichte und historisches Lernen. Festschrift für Ulrich Mayer zum 65. Geburtstag* (S. 242–255). Schwalbach/Ts.: Wochenschau.
- Schönemann, Bernd/Thünemann, Holger & Zülsdorf-Kersting, Meik (2011). *Was können Abiturienten?: Zugleich ein Beitrag zur Debatte über Kompetenzen und Standards im Fach Geschichte* (2. Aufl.). Münster: LIT-Verl.
- Schwabe, Astrid (2016). Multimedia statt Papier und Druckerschwärze – das «alte» Medium Zeitung in der digitalen Welt und der Geschichtsunterricht. In Christian Kuchler & Benjamin Städter (Hrsg.), *Zeitungen von gestern für das Lernen von morgen?: Historische Tagespresse im Geschichtsunterricht* (S. 161–177). Göttingen.
- Sochatzky, Florian, Schöner, Alexander, & Schreiber, Waltraud (Hrsg.) (2013). *Analyse von Schulbüchern als Grundlage empirischer Geschichtsdidaktik*. Stuttgart: Kohlhammer.
- Souvignier, Elmar & Dignath Ewijk, Charlotte van (2010). Pädagogische Interventionsforschung – ein historischer Rückblick. In Tina Hascher & Bernhard Schmitz (Hrsg.), *Pädagogische Interventionsforschung: Theoretische Grundlagen und empirisches Handlungswissen* (S. 12–30). Weinheim, München: Juventa.
- Stöber, Rudolf (2005). *Deutsche Pressegeschichte: Von den Anfängen bis zur Gegenwart* (2., überarb. Aufl.). Konstanz: UVK.
- Tenfelde, Klaus, & Urban, Thomas (Hrsg.). (2010). *Das Ruhrgebiet – Ein historisches Lesebuch: Zusammengestellt und bearbeitet von den Mitarbeiterinnen und Mitarbeitern des Instituts für soziale Bewegungen der Ruhr-Universität Bochum und der Stiftung Bibliothek des Ruhrgebiets*. Essen: Klartext.
- Waldis, Monika / Marti, Philipp & Nitsche, Martin (2015). Angehende Geschichtslehrpersonen schreiben Geschichte(n). *Zeitschrift für Geschichtsdidaktik*, 14(1), 63–86. <https://doi.org/10.13109/zfgd.2015.14e.1.63>
- Wineburg, Sam (2019). Warum historische Kompetenzen für die Auswertung von digitalen Quellen nicht ausreichend sind. In Sebastian Barsch, Andreas Lutter, & Christian Meyer-Heidemann (Hrsg.), *Fake und Filter: Historisches und politisches Lernen in Zeiten der Digitalität* (S. 105–120).