



## 50 Jahre Unterrichtswissenschaft – Themen der Forschung über das Lehren und Lernen im Wandel

Marcus Kubsch · Knut Neumann · Michael Rochnia · Cornelia Gräsel

Eingegangen: 21. September 2022 / Überarbeitet: 10. Januar 2023 / Angenommen: 13. Januar 2023 /  
Online publiziert: 14. Februar 2023  
© Der/die Autor(en) 2023

**Zusammenfassung** Seit 50 Jahren fokussiert die Zeitschrift *Unterrichtswissenschaft* auf die Erforschung von Fragen rund um das Lehren und Lernen, insbesondere in institutionalisierten Settings und im deutschsprachigen Raum. Der vorliegende Beitrag gibt einen Überblick darüber, welche Themen die Zeitschrift in den letzten 50 Jahren behandelt hat und wie sich die Themen im Verlauf der Entwicklung verändern. Es stellt sich die Frage, mit welchen Methoden man die historische Themenentwicklung erfasst. Eine Möglichkeit wäre, erfahrene Expert\*innen zu befragen, die die Entwicklung nicht nur lange verfolgt, sondern auch mitgestaltet haben. Die persönliche Expertise würde aber Gedächtnisverzerrungen und individuelle Bewertungen implizieren. Deshalb wurde im vorliegenden Beitrag auf das Verfahren der Computational Grounded Theory zurückgegriffen. Computational Grounded Theory basiert auf der Methode der Grounded Theory und nutzt KI-gestützte Verfahren, wie z. B. Natural Language Processing, um die subjektiven Teile des Interpretationsprozesses zu dokumentieren. Konkret wurde topic modeling genutzt, um aus den Abstracts sämtlicher in der 50-jährigen Geschichte der Unterrichtswissenschaft erschienenen Beiträge Themen zu destillieren. Die identifizierten Themen wurden durch die Autor\*innen mit den ihnen zugeordneten Beiträgen abgeglichen und gegebenenfalls adaptiert, bevor anschließend die (relative) Bedeutung der Themen im Verlauf der Geschichte untersucht wurde. Die Ergebnisse werden auch in Bezug zur Entwicklung der Lehr-Lernforschung der letzten 50 Jahre interpretiert.

---

✉ Marcus Kubsch · Knut Neumann  
IPN – Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik,  
Olshausenstraße 62, 24118 Kiel, Deutschland  
E-Mail: [kubsch@leibniz-ipn.de](mailto:kubsch@leibniz-ipn.de)

Michael Rochnia · Cornelia Gräsel  
Bergische Universität Wuppertal, Gaußstraße 20, 42119 Wuppertal, Deutschland

**Schlüsselwörter** Lehr-Lern-Forschung · Machine Learning · Natural Language Processing · Computational Grounded Theory · Wissenschaftsgeschichte

## 50 years of Unterrichtswissenschaft – topics of research on teaching and learning in transition

**Abstract** For 50 years, the journal *Unterrichtswissenschaft* has focused on researching issues related to teaching and learning, especially in institutionalized settings and in German-speaking countries. This article provides an overview of the topics the journal has covered over the past 50 years and how the topics have changed over the course of its development. The question arises, which methods can be used to capture the historical development of these topics? One possibility would be to interview experienced experts who have not only followed the development for a long time, but also helped to shape it. However, personal expertise would imply memory biases and individual evaluations. Therefore, the present paper resorted to the method of computational grounded theory. Computational grounded theory is based on the grounded theory method and uses AI-assisted procedures, such as natural language processing, to document the subjective parts of the interpretation process. Specifically, topic modeling was used to distill topics from the abstracts of all articles published in the 50-year history of *Unterrichtswissenschaft*. The identified themes were evaluated and adopted by the authors based on the articles that were most exemplary for each theme before subsequently examining the (relative) importance of the themes over the course of history. The results are interpreted in relation to the development of research on teaching and learning over the past 50 years.

**Keywords** Learning science · Machine learning · Natural language processing · Computational grounded theory · History of science

### 1 Einleitung

Unterricht ist eine spezifische Form pädagogischen Handelns. Das Besondere dabei ist, dass das pädagogische Handeln auf Lernen durch Lehren im Rahmen einer Institution abzielt (Kiel 2018). In der Regel kann Unterricht daher als ein Schnittpunkt von vier Aspekten verstanden werden: *Lernende* werden von einem *Lehrenden* in einer *Institution* zu einem bestimmten *Inhalt* unterrichtet. Die damit einhergehenden Phänomene besser zu verstehen, ist Anliegen der empirischen Unterrichtsforschung (vgl. Jank und Meyer 2002; Kiel 2018). Die (empirische) Unterrichtsforschung, auch als Lehr-Lernforschung bezeichnet<sup>1</sup>, geht grob drei Fragestellungen nach (Gräsel und Gniewosz 2011): Wie kann Lernen modelliert werden? Welche Wirkungen

---

<sup>1</sup> Manchmal wird die Unterrichtsforschung dadurch von der Lehr-Lernforschung abgegrenzt, dass sie auf das Lehren und Lernen in schulischen Institutionen fokussiert (vgl. Gräsel und Gniewosz 2011); dieser Begriffsverwendung folgt der vorliegende Beitrag. Unterrichts- bzw. Lehr-Lernforschung arbeiten empirisch, daher wird dieses Adjektiv manchmal für ihre Bezeichnung verwendet („empirische Unterrichtsforschung“). Das impliziert aber keinen Bedeutungsunterschied. Da die Zeitschrift *Unterrichtswissenschaft*

haben unterschiedliche Formen des Lehrens? Welche (weiteren) individuellen und institutionellen Merkmale haben Einfluss auf das Lehren und Lernen? Diese Fragen wurden und werden unter anderem in der Zeitschrift *Unterrichtswissenschaft – Zeitschrift für Lernforschung*<sup>2</sup> verhandelt.

Seit nunmehr 50 Jahren fokussiert die *Unterrichtswissenschaft* auf die Erforschung von Fragen rund um das (unterrichtliche) Lehren und Lernen, um ein besseres Verständnis entsprechender Phänomene zu erreichen. Die erste Ausgabe der Zeitschrift erschien 1973; hervorgegangen ist sie aus der *Zeitschrift für Programmierendes Lernen*. Seitdem erschienen anfangs zwei bis drei, seit einiger Zeit regelmäßig vier Hefte pro Jahr mit jeweils vier bis acht Beiträgen. Das sind pro Jahr ungefähr dreißig oder, über 50 Jahre hinweg, insgesamt ungefähr 1200 Beiträge. Deren Spektrum an Textsorten reicht von theoretischen Beiträgen wie Synthesen des Forschungsstands über Positionspapiere bis hin zu Beiträgen, die empirische Studien darstellen. Sie alle decken ein breites Spektrum an Themen aus den Bereichen Vorschule, Schule, Studium, Beruf sowie Freizeit ab. Die Veröffentlichungen der *Unterrichtswissenschaft* haben ohne Frage zu einer substanziellen Weiterentwicklung der Lehr-Lernforschung beigetragen und neben einem besseren Verständnis des (unterrichtlichen) Lehrens und Lernens auch zu einer Verbesserung desselben beigetragen. Aber was waren konkret die Themen der letzten 50 Jahre und wie haben sich diese über die Zeit hinweg entwickelt?

Im Vorfeld dieses Beitrags haben wir uns zunächst mit den vorliegenden Veröffentlichungen der letzten 50 Jahre beschäftigt und versucht, die groben Entwicklungslinien, die sich in der Zeitschrift *Unterrichtswissenschaft* finden lassen, zu identifizieren. Dies haben wir auch im Kreis der Herausgeber\*innen diskutiert. Hier das Ergebnis unserer Elaborationen: In den letzten Jahren scheinen für die Lehr-Lernforschung Arbeiten zur Unterrichtsqualität in der Schule wichtig gewesen zu sein, deren differenzierte Analysen unterrichtlicher Prozesse in Relation zum Lernerfolg die drei so genannten Basisdimensionen des Unterrichts hervorgebracht haben (Praetorius und Nehring 2020). Ferner entstammten viele Aufsätze der Kompetenzforschung, die sich intensiv mit der Frage der Ziele unterrichtlichen Handelns auseinandergesetzt hat und neben einer stärker domänenspezifischen Orientierung der Unterrichtsforschung vor allem auch die Erkenntnis gebracht hat, dass nicht nur die unterrichtlichen Prozesse, sondern auch die Ergebnisse des unterrichtlichen Lehrens und Lernens multidimensional sind (Zlatkin-Troitschanskaia et al. 2016). Nicht zu vergessen der große Bereich der Forschung zu Lehrkräften und ihrer Professionalisierung, der in den letzten Jahren einen starken Aufschwung erlebte. Diese Forschung hat wesentlich dazu beigetragen, die Lehrkraft als zentrales Element des Unterrichts und die Entwicklung der professionellen Handlungskompetenz von Lehrkräften als wichtige Säule qualitativ hochwertigen Unterrichts zu etablieren (Kleickmann und Hardy 2019). Welche Themen in der Zeitschrift von Bedeutung waren, war natürlich einem zeitlichen Wandel unterworfen. In den 1970er-Jahren war offenbar die Medienforschung ein wichtiges Thema. In den 1980er-Jahren wie-

---

trotz ihres Titels das Lehren und Lernen in verschiedenen Institutionen thematisiert, wird im Folgenden von Lehr-Lernforschung gesprochen, wenn nicht Forschung zu schulischem Unterricht adressiert wird.

<sup>2</sup> Im Folgenden: *Unterrichtswissenschaft*.

derum scheinen unter dem Eindruck des vorangegangenen Jahrzehnts und der kognitiven Wende die Selbstständigkeit und die Lernenden in ihrer Individualität im Vordergrund gestanden zu haben und Fragen der Selbstregulation, Motivation und weiterer affektiver Aspekte des unterrichtlichen Lehrens und Lernens stärker in den Vordergrund gerückt zu sein. Die 1990 Jahre, insbesondere die zweite Hälfte, waren geprägt durch die empirische Wende – insbesondere in den Fachdidaktiken der Mathematik und der Naturwissenschaften. Bedingt dadurch rückte die Erforschung unterrichtlicher Prozesse des Lehrens und Lernens in diesen Fächern, aber auch in anderen, in den Vordergrund. Auch die Medienforschung erfuhr in den 1990er-Jahren eine Renaissance, nun mit Fokus auf Multimedia und computerbasiertem Lernen.

Aber waren dies wirklich die großen Themen der Lehr-Lernforschung? Ist ihre Entwicklung tatsächlich so verlaufen? Was ist mit Themen wie der Moralerziehung? Und waren nicht auch schon in den Jahren vor dem Jahr 2000 die Ergebnisse von Lehrkräfteaus- und -weiterbildung von Relevanz? Eine andere Autor\*innengruppe hätte wahrscheinlich andere Themen benannt oder würde die Entwicklung der Themen anders einschätzen. Schließlich stellt sich die Frage, ob diese wirklich im eigentlichen Sinne abgrenzbare Themen sind, also Inhalte, die von mehreren Forschungsgruppen in unterschiedlicher Weise bearbeitet wurden, oder möglicherweise nur einzelne, wengleich wichtige und sicher auch bedeutsame Beiträge zur Lehr-Lernforschung. Ähnlich schwer einzuschätzen sind aktuelle Themen und inwieweit sie eine eigenständige Entwicklung begründen oder möglicherweise zur Veränderung bestehender Themen beitragen. Hier wäre z. B. das im Kontext Bildung zunehmend bedeutsame Thema der künstlichen Intelligenz (KI) zu nennen, das sicher an Bedeutung zunehmen wird. Dieses könnte z. B. die Lehrkräfteaus- und -weiterbildung maßgeblich beeinflussen, z. B. in Richtung eines stärkeren Fokus auf die Orchestrierung KI-gestützten Unterrichts. Die Fragen, die sich anhand dieser Überlegungen stellen sind:

1. Welche Themen sind über 50 Jahre in der *Unterrichtswissenschaft* Themen der deutschsprachigen empirischen Lehr-Lernforschung gewesen?
2. Wie haben sich diese Themen und damit dieser Bereich der empirischen Lehr-Lernforschung über die Zeit verändert?

Der vorliegende Beitrag geht dieser Frage mit Hilfe eines Verfahrens nach, das einen textanalytischen Zugang mit der Systematik moderner KI-gestützter Verfahren verknüpft: *computational grounded theory* (Nelson 2020). Computational Grounded Theory, basiert auf der Methode der Grounded Theory und nutzt KI-gestützte Verfahren, um die subjektiven Teile des Interpretationsprozesses zu dokumentieren. Im vorliegenden Beitrag wurden zunächst aus den Abstracts sämtlicher in der 50-jährigen Geschichte der Unterrichtswissenschaft erschienenen Beiträge mit Hilfe eines solchen Verfahrens einschlägige Themen destilliert. Diese wurden durch die Autor\*innen mit den ihnen zugeordneten Beiträgen abgeglichen und gegebenenfalls adaptiert, bevor anschließend die (relative) Bedeutung der Themen im Verlauf der Geschichte analysiert wurde. Im Folgenden gehen wir zunächst kurz auf die Bedeutung von Texten als Forschungsdaten sowie (traditionelle) Verfahren der Textanalyse

ein, bevor wir unsere methodische Vorgehensweise näher beschreiben und die Ergebnisse unserer Analysen präsentieren. Abschließend diskutieren wir einerseits unsere inhaltlichen Erkenntnisse und reflektieren andererseits die Möglichkeiten und Grenzen der verwendeten Methodik und ihr Potenzial für Textanalysen in der empirischen Unterrichts- bzw. Bildungsforschung.

## 2 Theoretischer Hintergrund

Texte spielen als Forschungsdaten eine wichtige Rolle in der Bildungsforschung. Textdaten umfassen dabei ein breites Spektrum – von Lernendenantworten in Freitextaufgaben (z. B. Rosenberg und Krist 2020) über Schulbücher (z. B. Christianson et al. 2020) und Bildungsstandards (Shaw et al. 2016) bis hin zu Forschungsarbeiten (Odden et al. 2021). Die (teils extreme) inhaltliche Dichte und interpretatorische Offenheit von Textdaten führt dabei zu ganz eigenen Herausforderungen für den Erkenntnisgewinnungsprozess aus Textdaten. Für die Analyse von Texten lassen sich verschiedene forschungsmethodische Zugänge unterscheiden. In den Geistes- und Kulturwissenschaften – auch in den entsprechenden Didaktiken – spielen historische und/oder hermeneutische Ansätze eine wichtige Rolle, in denen das Verstehen und Interpretieren von Texten zentral ist (z. B. Schröder 2012; vgl. Rittelmeyer et al. 2001). In der empirischen Lehr-Lernforschung kann man bei der Analyse von Textdaten allgemein zwischen qualitativen (z. B. Glaser und Strauss 2017; Kuckartz und McWhertor 2014; Mayring 2014) und quantitativen Verfahren unterscheiden (z. B. Blei 2012; Landauer 2014).

Qualitative Verfahren beruhen dabei typischerweise auf einer interpretativen Epistemik, die die Realität und ihre Bedeutung als Produkt einer sozialen und kontext-abhängigen Co-Konstruktion versteht. Qualitative Verfahren können so zu tiefen und reichhaltigen Beschreibungen von Phänomenen führen und insbesondere die Entwicklung von Theorien unterstützen. Eine Herausforderung bei der Anwendung qualitativer Verfahren stellt der hohe Arbeitsaufwand dar. So müssen umfangreiche Textmengen wiederholt gelesen, interpretiert und gegebenenfalls kategorisiert werden. Der große Aufwand für die qualitative Analyse größerer Textkorpora kann im Rahmen von gut finanzierten, langfristigen Forschungsprojekten geleistet werden. Für einen einleitenden Aufsatz unseres Jubiläumsheftes hätte sich schwerlich ein Geldgeber, oder eine Geldgeberin, finden lassen, der die qualitative Analyse von 50 Jahrgängen der *Unterrichtswissenschaft* finanziert hätte (und das zu Recht, denn so wichtig ist die Fragestellung nun auch nicht). Generell ist bei qualitativen Verfahren also abzuwägen, welcher Umfang an Textdaten sich für entsprechende Analysen im Verhältnis zum Erkenntnisgewinn eignet. Ferner ergibt sich die epistemische Limitation, dass Phänomene, die erst in großen, unstrukturierten Textmengen sichtbar werden, den Verfahren nicht zugänglich sind (Bail 2014). Eine weitere Limitation qualitativer Verfahren kann sein, dass ihre Ergebnisse schwierig zu validieren und zu reproduzieren sind, da die subjektiven Urteile der Forschenden Teil der Analyse sein können (Biernacki 2012).

Quantitative (computationale) Verfahren der Textanalyse sind im Gegensatz zu qualitativen Verfahren eher reproduzierbar (z. B. aufgrund der Dokumentation von

Analyseskripten). Zudem liegen für Entscheidungen im Analyseprozess häufig objektive(re) Kriterien vor. Nicht zuletzt erlauben quantitative Verfahren durch ihre exakten Berechnungsvorschriften, insbesondere bei Verwendung von Computern, die Verarbeitung großer Textmengen. Diesen Vorteilen quantitativer Verfahren steht das Problem gegenüber, dass die Ergebnisse dieser Verfahren, z.B. die Zahl der Wörter oder die syntaktische Nähe von Wörtern in einem Text zueinander, nicht direkt interpretierbar sind (Blei 2012). Folglich ist immer eine inhaltliche Interpretation der Ergebnisse durch die Forschenden nötig. Hierbei besteht das Risiko, dass aufgrund der hohen technischen Komplexität der Verfahren die Ergebnisse fehlinterpretiert werden (oder bereits die Methode unangemessen verwendet wird, Baumer et al. 2017). Des Weiteren können durch die reduzierte Darstellung der Ergebnisse quantitativer Methoden wichtige, kontextuelle Faktoren verborgen bleiben oder nicht vorhandene Zusammenhänge suggeriert werden (vgl. Rost et al. 2013).

Mit der *Computational Grounded Theory* stellt Nelson (2020) ein Verfahren vor, das quantitative und qualitative Verfahren kombiniert, mit dem Ziel, die individuellen Stärken zu kombinieren und Schwächen zu beheben. Die *Computational Grounded Theory* nach Nelson (2020) umfasst drei Schritte. Im ersten Schritt (*pattern detection*) werden mit Hilfe quantitativer (computational) Verfahren potenziell umfangreiche Textdaten auf zentrale Merkmale reduziert – zum Beispiel in Form von geordneten Wortlisten, Wortnetzwerken oder Textgruppen. Der Vorteil der Nutzung quantitativer (computational) Verfahren ist dabei, dass eine große Menge (unstrukturierter) Textdaten in einem reproduzierbaren Verfahren einer menschlichen Interpretation zugänglich wird. Im zweiten Schritt (*pattern refinement*) soll durch ein „computationally guided deep reading of the text“ (Nelson 2020, S. 6) die Plausibilität der vom Computer gefundenen Muster in den Daten unter Nutzung der interpretativen Möglichkeiten des Menschen überprüft werden, um eine inhaltliche Deutung der gefundenen Muster vorzunehmen. Diese ersten beiden Schritte sind dabei iterativ zu verstehen, d. h. die Ergebnisse eines *pattern refinement* können zu einer erneuten *pattern detection* und damit zur schrittweisen Optimierung der Interpretation führen. Schließlich wird – wenn möglich – in einem dritten Schritt, der *pattern confirmation*, die Generalisierbarkeit der zuvor erarbeiteten Ergebnisse überprüft. Dieser Schritt dient also der Prüfung, inwieweit die identifizierten Muster von Merkmalen und ihre Interpretation nicht nur Artefakt eines quantitativen (computational) Verfahrens oder einer menschlichen Fehlinterpretation sind. Dazu soll analysiert werden, ob die Muster (und Interpretationen) über die sämtlichen Textdaten hinweg wiederholt identifizierbar sind.

Die Methode der *Computational Grounded Theory* stellt eine nahezu ideale Methodik zur Beantwortung der eingangs formulierten Forschungsfragen dar. Insbesondere kann die Kombination aus computergestützter Identifikation möglicher Muster in den Texten, kombiniert mit menschlicher Interpretation helfen, Verzerrungen durch menschliche Vorerfahrungen einerseits und wenig informative Statistiken zu Textmerkmalen andererseits zu vermeiden.

### 3 Methode

Die *Unterrichtswissenschaft* ist als Zeitschrift für empirische Lehr-Lernforschung eine Zeitschrift mit einem seit 50 Jahren relativ konstanten Scope. Dies bedeutet, dass die Zeitschrift sich in besonderer Weise eignet, die Strömungen der deutschsprachigen Lehr-Lernforschung im letzten halben Jahrhundert zu analysieren. Damit können die Entwicklung der Zeitschrift und damit bis zu einem bestimmten Grad auch die Entwicklung der deutschsprachigen Lehr-Lernforschung nachgezeichnet werden. Um den eingangs formulierten Fragestellungen nach den zentralen Themen der deutschsprachigen empirischen Lehr-Lernforschung und deren Entwicklung zu beantworten, wurden die Abstracts sämtlicher in der 50-jährigen Geschichte der *Unterrichtswissenschaft* erschienen Beiträge mittels *computational grounded theory* (Nelson 2020) analysiert. Die Abstracts wurden verwendet, da die Zeitschrift nicht so vollständig digitalisiert vorliegt, dass sie komplett ausgewertet werden kann. Da die Abstracts die zentralen Inhalte der Aufsätze zusammenfassen und dazu auch die entsprechenden Kernbegriffe verwenden, kann aber davon ausgegangen werden, dass sie die Gesamtaufsätze gut repräsentieren. Für die Analyse wurden zunächst alle Abstracts der Beiträge in der *Unterrichtswissenschaft* digitalisiert. Abstracts, die nicht bereits digital vorlagen, wurden gescannt.

Im ersten Schritt der Analyse, der *pattern detection*, wurden mit Hilfe eines KI-gestützten Verfahrens Themen im Sinne von Mustern von Textmerkmalen identifiziert, die im zweiten Schritt, dem *pattern refinement*, durch die Autoren in einem iterativen Verfahren charakterisiert wurden. Da wir mit dem kompletten Textkorpus der Abstracts der *Unterrichtswissenschaft* arbeiteten, stellt sich die Frage nach der Generalisierbarkeit der Ergebnisse nur eingeschränkt, so dass der dritte Schritt der *pattern confirmation* nicht angewandt wurde. Wir verstehen jedoch die durchgeführten Analysen zur Entwicklung der Themen zumindest teilweise als eine Prüfung der Konsistenz unserer Interpretation der identifizierten Themen. In den folgenden Abschnitten gehen wir detaillierter auf die zentralen Aspekte unseres methodischen Vorgehens ein: Wir beschreiben zunächst die unseren Analysen zugrundeliegende Daten, dann die durchgeführten Schritte der Aufbereitung der Daten für eine quantitative (computationale) Analyse und gehen zuletzt auf die einzelnen Schritte der Analyse selbst ein.

#### 3.1 Datenbasis

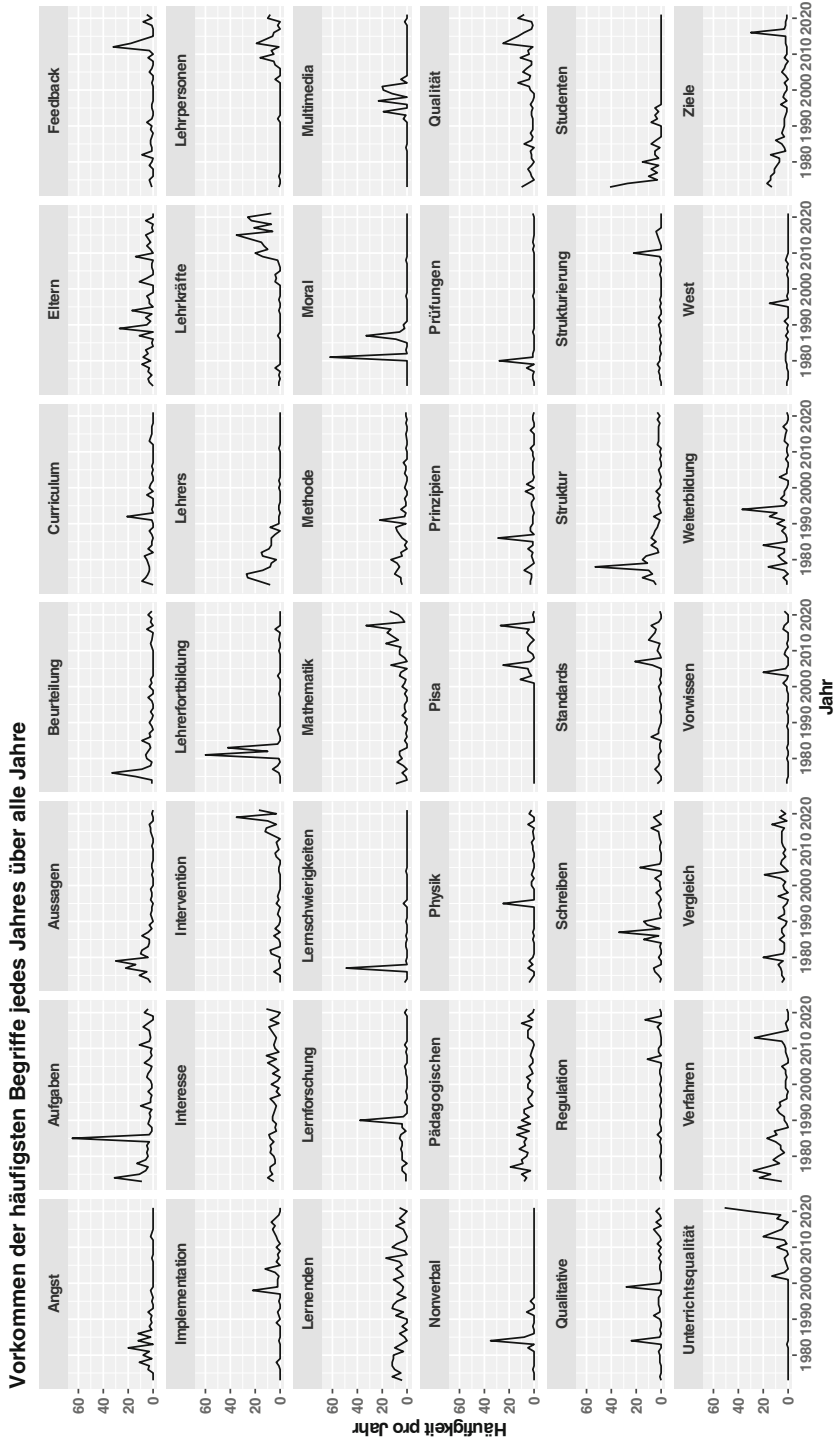
Mit Blick auf das Ziel der Untersuchung, war eine Datenbasis notwendig, die einen möglichst großen Zeitraum abdeckt. Konkret wurden möglichst viele Texte aus möglichst vielen Jahren benötigt. Die *Unterrichtswissenschaft* bietet in diesem Zusammenhang eine ideale Datenquelle, da sie seit nunmehr 50 Jahren jährlich ca. 25 Artikel in vier Heften veröffentlicht und die einzige deutschsprachige Zeitschrift ist, die sich dezidiert der „Lehr-Lernforschung“ widmet, und zwar sowohl unter fachspezifischer als auch fachübergreifender Perspektive. Damit bietet die Zeitschrift eine 50-jährige ununterbrochene Kette von Artikeln, die nicht nur die Themen der einzelnen Jahre, sondern auch die Entwicklung der Themen über die Jahre abbilden (können).

Als Datenbasis wurden zunächst alle in der Zeitschrift erschienenen Artikel gesammelt. Dies geschah im Frühjahr 2022. Die Jahrgänge ab 2018 bis in die Gegenwart sind online auf der Homepage der Zeitschrift verfügbar. Ältere Jahrgänge mussten aus dem Präsenzbestand der Bibliotheken der Bergischen Universität Wuppertal und der Christian-Albrechts-Universität zu Kiel beschafft werden. Studentische Hilfskräfte beider Universitäten scannten die Jahrgänge 1973 bis 2017 und speicherten diese als PDF ab. Um den Arbeitsaufwand im Rahmen zu halten, wurde auf die Abstracts der Originalartikel als summative Inhaltsangabe fokussiert. Falls kein Abstract verfügbar war (dies war bei älteren Jahrgängen der Fall), wurde die erste Seite des Artikels gescannt. Hierbei kam es vor, dass durch die Anordnung der Beiträge in der Zeitschrift und den verwendeten Scannern Teile des Literaturverzeichnisses des vorangegangenen Artikels mitgescannt und in die Analyse einbezogen wurden. Es wurden theoretische und empirische Beiträge in den Sprachen Deutsch und Englisch berücksichtigt. Buchbesprechungen oder anderweitige nicht wissenschaftliche Inhalte wurden aus der Analyse ausgeschlossen. Insgesamt lagen am Ende der Materialsammlung 1218 Abstracts als PDF-Dateien beginnend mit dem ersten Heft 1973 und endend mit Heft 4 des Jahres 2021.

### 3.2 Datenaufbereitung

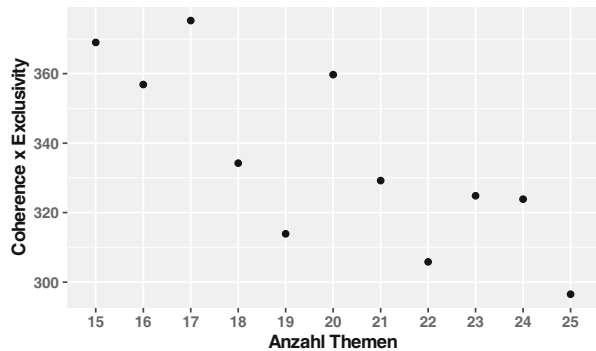
Die Datenaufbereitung erfolgte in einem teils iterativen Verfahren unter Verwendung des R-Pakets *Quanteda* (Benoit et al. 2018). Ziel war es, inhaltlich nicht relevanten Umfang bzw. Komplexität in der Datenbasis zu reduzieren und die Daten einer inhaltlichen Analyse zugänglicher zu machen. Es umfasste die folgenden Schritte: 1) Zunächst wurden Wörter deutscher und englischer Sprache ohne semantische Bedeutung – so genannte *stopwords* – aus dem Textkorpus entfernt. Hierfür wurden die in *Quanteda* enthaltenen *stopword-List*en *snowball* und *stopwords-iso* genutzt. Beispiele für stopwords sind zum Beispiel *in*, *aber*, *und*, *auf*, *der*, *die*, *das*, *are*, *a*, *about*, *too*, *then* etc. 2) Anschließend wurden Worthäufigkeiten pro Jahr berechnet und inhaltsleere Wörter mit hoher Frequenz der *stopword-Liste* hinzugefügt – zum Beispiel Begriffe wie *Abstract*, *Herausgeber*, *Zusammenfassung*, *GmbH*, *Thema*, *Zeitschrift*, etc. Auch häufig vorkommende Namen wie *Blömeke*, *Kunter*, etc. wurden der *stopword-Liste* hinzugefügt. Dieser Schritt wurde iteriert, bis die zehn häufigsten Wörter pro Jahr als inhaltlich zu werten waren – z.B. *Unterricht*, *Lernen*, *Schülerinnen*, etc. 3) Um die Textmenge weiter zu reduzieren und eine spätere inhaltliche Interpretation zu erleichtern, wurden nun automatisiert Wörter entfernt, welche entweder sehr häufig oder sehr selten vorkamen. Die zugrundeliegende Idee bei diesem Schritt ist, dass sehr häufige Wörter zeitlich invariant sind und spezifischere Wörter inhaltliche Veränderungen anzeigen. (So wäre es keine Erkenntnis, wenn man herausfinden würde, dass *Lernen* und *Wissen* große Themen in der Unterrichtswissenschaft sind). Basierend auf Erfahrungswerten und Anhaltspunkten in der Literatur (Odden et al. 2020, 2021) wurde entschieden, Wörter, die in mehr als 50 % der Artikel vorkamen, und Wörter, die in nur einem Artikel vorkamen, aus der Datenbasis zu entfernen. Der Grenzwert 50 % wurde analog zu Schritt 2) iterativ durch Inspektion der zehn häufigsten Wörter pro Jahr festgelegt. Der Grenzwert von 50 % ist dabei für computerbasierte Textanalysen angemessen, in denen zwischen





**Abb. 1** Häufigste Begriffe jedes Jahres über Erscheinungsjahre der Unterrichtswissenschaft

**Abb. 2** Produkt aus coherence und exclusivity für 15 bis 25 Themen



25 % und 99 % als Grenzwert nicht ungewöhnlich sind (Odden et al. 2021). Abb. 1 zeigt für die resultierende Datenbasis die Häufigkeit des häufigsten Wortes pro Jahr für alle Jahre.

### 3.3 Datenanalyse

Die Analyse mittels *Computational Grounded Theory* (Nelson 2020) verlief in drei Schritten: Es wurden 1.) mit Hilfe computergestützter Datenanalyse Themen im Textkorpus identifiziert und 2.) durch eine inhaltliche Analyse der den Themen zugeordneten Texte charakterisiert. Abschließend wurden 3.) die Themen in ihrer zeitlichen Entwicklung analysiert.

#### 3.3.1 Identifikation der Themen

Um Themen in Texten zu identifizieren, wurde das so genannte *topic modeling* zur Identifikation von Mustern von Textmerkmalen eingesetzt (z. B. Blei 2012; Blei und Lafferty 2007). Muster von Textmerkmalen bedeutet, dass bestimmte Wörter besonders häufig und in Kombination mit anderen Wörtern vorkommen. Beispielsweise kommt das Wort *Weiterbildung* häufig in Kombination mit *Erwachsenenbildung* vor. Beim *topic modeling* muss die Anzahl der zu schätzenden Themen vorgegeben werden. In Folge werden, ähnlich wie bei anderen explorativen Verfahren, wie zum Beispiel der latenten Profilanalyse, zunächst viele Modelle berechnet und dann anhand von Metriken ein Modell mit einer bestimmten Anzahl von Themen ausgewählt. Dabei muss zwischen einfacher Repräsentation (eher wenige Themen) und Distinktion (eher viele Themen) abgewogen werden. Mehrere hundert Themen würden viele Strömungen der Lehr-Lernforschung unterscheiden, wären also schwer darstellbar. Zehn Themen wären leicht darstellbar, könnten jedoch bestimmte Strömungen der Lehr-Lernforschung unterschlagen. Es gilt daher, eine optimale Balance zwischen Interpretierbarkeit und Distinktion zu finden. Basierend auf ähnlichen Arbeiten (z. B. Odden et al. 2021) haben wir zunächst mit Hilfe des R-Pakets *stm* (Roberts et al. 2019) Modelle mit 1 bis 50 Themen geschätzt und die Metriken *coherence* (Mimno et al. 2011) und *exclusivity* berechnet. Eine hohe *coherence* liegt dann vor, wenn Wörter, die für ein Thema am wahrscheinlichsten sind, häufig zusammen auftreten. Dieses Maß korreliert sehr gut mit menschlichen Einschätzungen der Qualität

von Themen. Allerdings kann dieses Maß einfach maximiert werden, wenn einige wenige Themen von sehr häufigen Wörtern dominiert werden. *Exclusivity* ist umso höher, je weniger ein Wort, das für ein Thema sehr wahrscheinlich ist, auch in anderen Themen vorkommt. Insofern stellt *exclusivity* ein Gegengewicht zu *coherence* dar. Es wird empfohlen, die Anzahl von Themen basierend auf beiden Kriterien auszuwählen (Roberts et al. 2019). Für die geschätzten Modelle war das Produkt von *coherence* und *exclusivity* für die Zahl von 17 Themen maximal (Abb. 2). Deshalb wurde nach einem positiven *face validity check* (Odden et al. 2021) das Modell mit 17 Themen für die weitere Analyse ausgewählt.

### 3.3.2 Charakterisierung der Themen

Zur weiteren Beschreibung der Themen wurden für jedes der 17 Themen die fünf Texte (d. h. Abstracts) herangezogen, die den jeweiligen Themen mit größter Wahrscheinlichkeit zuzuordnen waren. Anhand der (aus dem vorherigen Schritt stammenden) häufigsten mit einem Thema assoziierten Wörter und der in diesem Schritt identifizierten Texte wurde zu jedem Thema von einem der Autoren dieses Beitrags eine zusammenfassende Beschreibung (im Sinne eines Memos) formuliert. Jede Beschreibung wurde anschließend von einem anderen Autor, wiederum auf Basis der häufigsten mit einem Text assoziierten Wörter und Texte, kritisch geprüft und überarbeitet. Die überarbeiteten Beschreibungen wurden abschließend noch einmal im Autorenkreis diskutiert, bis ein finaler Konsens über die Beschreibung hergestellt war.

### 3.3.3 Analyse der zeitlichen Entwicklung

Um die zeitliche Entwicklung der Themen zu untersuchen, wurde das *topic modeling* erneut durchgeführt, diesmal mit dem Erscheinungsjahr des jeweiligen Texts als Kovariate. Dadurch kann für jedes Thema der Anteil an den Texten eines Jahres bestimmt und analysiert werden, wie sich dieser Anteil über die Zeit entwickelt.

## 4 Ergebnisse

### 4.1 Forschungsfrage 1: Welche Themen sind über 50 Jahre Unterrichtswissenschaft Themen der empirischen Lehr-Lernforschung gewesen?

Abb. 3 führt für jedes der 17 Themen die fünf Wörter mit den größten Wahrscheinlichkeiten, in dem entsprechenden Thema vorzukommen ( $\beta$ ) an. Je höher diese Wahrscheinlichkeit für ein Wort in einem Thema ist, desto häufiger wird es in Texten des entsprechenden Themas vorkommen. Folglich sind die Wörter mit dem größten  $\beta$  am prägendsten für die inhaltliche Deutung eines Themas. Es fällt auf, dass sich einige Themen inhaltlich intuitiv deuten lassen, während dies bei anderen schwieriger ist. So ist bei Thema 7 – Professionelle Kompetenzen von Lehrkräften ein Bezug zu



**Abb. 3** Häufigste Begriffe pro Thema und deren jeweilige Wahrscheinlichkeit in dem entsprechenden Thema ( $\beta$ )

professioneller Kompetenz von Lehrkräften klar erkennbar, wohingegen Thema 12 – Förderung domänenspezifischen Wissens thematisch breiter erscheint.

Dieser Unschärfe begegnet nun die im Folgenden dargestellte inhaltliche Analyse auf Basis der zu den Themen zugehörigen Texte. (Die Themen sind durch Nummern und das jeweils häufigste Wort bezeichnet).

**Thema 1 – Artefakt** Thema 1 ist kein inhaltlich geprägtes Thema, sondern identifiziert Text mit vielen Zitationen, Inhaltsverzeichnissen, und Quellenverzeichnissen. Dieses Thema ist v. a. durch die Scans, welche z. T. Literaturverzeichnisse vorangegangener Artikel beinhalten, zu erklären. Hierzu passt auch, dass dieses Thema praktisch verschwindet, sobald die originär digital vorliegenden Versionen der Artikel für die Analysen verwendet wurden. Das Thema wurde von folgenden Analysen ausgeschlossen.

**Thema 2 – Schreiben und Lernen** Der Fokus von Thema 2 liegt auf dem Zusammenhang von Schreiben und Lernen. Die Beziehung zwischen Schreiben und Lernen ist dabei doppelt bestimmt: Wissen beeinflusst die Textproduktion und die Textproduktion reorganisiert wiederum das Wissen. Auch das Thema funktionaler Analphabetismus gehört zu diesem Thema. Es wird kritisch betrachtet, dass Menschen, die an funktionalem Analphabetismus leiden, kaum an Schriftkultur partizipieren können. Potenzielle Fördermöglichkeiten werden diskutiert.

**Thema 3 – Fort- und Weiterbildung** In Thema 3 liegt der Fokus auf der Fort- und Weiterbildung von Erwachsenen. Häufig wird dabei auch die Erfassung von Kompetenzen in verschiedenen Domänen thematisiert. Auch die Fortbildung von Lehrkräften aller Schulformen ist Teil dieses Themas. Die Arbeiten zeigen eine große Bandbreite: Die Beiträge reichen von der Evaluation spezieller Angebote über Positionspapiere bis hin zu Erfahrungsberichten. Auch die Kommunikation von Fort- und Weiterbildungsangeboten an potenzielle Teilnehmer\*innen wird in den Blick genommen.

**Thema 4 – Erfassung von Unterrichtsqualität und ihren Korrelaten** Thema 4 fokussiert auf Arbeiten zur Erfassung von Unterrichtsqualität und Zusammenhängen von Unterrichtsqualität mit weiteren Maßen – zentral ist dabei der schulische Unterricht. Hierbei wird die Erfassung von Kompetenzen der Schüler\*innen in verschiedenen Domänen berücksichtigt und hinsichtlich ihrer unterrichtlichen Relevanz diskutiert. Dabei stehen empirische Arbeiten im Fokus, welche verschiedene Fächer und Alterskohorten abdecken. Beispielsweise wird die naturwissenschaftliche Bildung im Kindergarten, aber auch die Lehrkräftebildung beforscht, die für qualitätsvollen Unterricht eine Voraussetzung ist. Auch werden verschiedene Methoden, Unterrichtsqualität zu erfassen, kritisch beleuchtet. Im Allgemeinen diskutiert Thema 4, überwiegend auf empirischer Basis, was qualitativ hochwertigen Unterricht ausmacht.

**Thema 5 – Lernstrategien** Das fünfte Thema widmet sich dem Erwerb und der Förderung von Lernstrategien. Im Zuge dessen sollen elaborative, reduktive und metakognitive Prozesse von Lernenden gefördert werden. Es werden Lernstrategien beschrieben und hinsichtlich ihres individuellen Profits für die Lernenden diskutiert. Die in Thema 5 verorteten Arbeiten zu Lernstrategien zeichnen sich durch einen kognitionspsychologischen Schwerpunkt aus, in dem Lernen als Prozess der Informationsverarbeitung verstanden wird. Die Mehrheit der Beiträge des fünften Themas instruiert Lernende theoriegeleitet und prüft die gegebenen Instruktionen empirisch hinsichtlich ihrer Wirkung. Übergeordnetes Ziel ist dabei die Analyse von Unterricht unter dem Gesichtspunkt der Sequenzierung des Lehrstoffs.

**Thema 6 – Fremdsprachenunterricht** Thema 6 umfasst verschiedene Arbeiten zum Fremdsprachenunterricht. Die Arbeiten adressieren Fremdsprachenunterricht auf mehreren Ebenen. So werden Fragen zum Kommunikationsbegriff und zu Kommunikationsstrukturen oder etwas, dass das Selbstverständnis der Fremdsprachendidaktik im Kontext der Einbettung des Fremdsprachenunterrichts in schulische Strukturen betrifft, abgehandelt; auch Arbeiten zu Fragen von Leistungen in und Einstellungen zum Fremdsprachenunterricht kommen vor. Methodisch liegt eine Mischung aus theoretischen Arbeiten, die aus dem Forschungsstand Konzepte des Fremdsprachenunterrichts begründen und experimentellen Arbeiten u. a. zur Einstellung von Erwachsenen gegenüber Englisch als Zielsprache vor.

**Thema 7 – Professionelle Kompetenzen von Lehrkräften** Thema 7 fokussiert auf professionelle Kompetenzen von Lehrkräften als Voraussetzung für erfolgreiches Unterrichtshandeln. Im Vordergrund stehen dabei fachübergreifende und bildungswissenschaftliche Aspekte professioneller Kompetenz und deren Bedeutung für das Lernen der Schüler\*innen. Die Rolle subjektiver Theorien als Voraussetzung für erfolgreiches Handeln und die Frage nach der Eignung von Lehrkräften und deren Feststellung sind ebenfalls Schwerpunkte des siebten Themas. In der Summe fokussiert Thema 7 auf die professionelle Kompetenz bzw. das professionelle Handeln von Lehrkräften im Unterricht, dessen Determinanten und wie sich entsprechende Kompetenz bzw. entsprechendes Handeln feststellen lässt. Methodisch handelt es sich mehrheitlich um theoretische bzw. Positionspapiere.

**Thema 8 – Kompetenzen in Mathematik und Naturwissenschaften** Thema 8 fokussiert auf die schulisch erworbenen Kompetenzen von Schüler\*innen bzw. die entsprechenden Leistungen in Mathematik und Naturwissenschaften und deren Determinanten. Unter anderem fallen hierunter die Berichte zu Ergebnissen eines groß angelegten Fortbildungsprogramms für Lehrkräfte in Hamburg – u. a. eine Studie zur Bedeutung kompetenziellen Feedbacks im Fach Biologie oder eine Studie zu schulisch erworbenen Kompetenzen in Mathematik im Vergleich zwischen beruflichen und allgemeinbildenden Gymnasien. Der überwiegende Teil der Arbeiten ist empirischer Natur.

**Thema 9 – Sozioökonomische und geschlechtsspezifische Faktoren** Gegenstand von Thema 9 ist die Beurteilung von Schüler\*innen durch Lehrkräfte. Dabei geht es um die vielen Facetten der Beurteilung: So können sozioökonomische Faktoren – wie das Bildungsniveau der Eltern – die Beurteilung von Schüler\*innen verzerren. Eine ähnliche Rolle kann das Geschlecht spielen, beispielsweise bei der Beurteilung mathematischer und sprachlicher Begabung. Insgesamt adressiert das neunte Thema Faktoren, welche die Objektivität, Reliabilität und Validität von Leistungsbeurteilung einschränken. Zuletzt geht es auch um die Beurteilung der Lehrkräfte, u. a. im Rahmen von Lehrproben und die Güte entsprechender Urteile.

**Thema 10 – Textverstehen und Textverständlichkeit** Thema 10 bezieht sich auf das Lernen aus Texten. Dies schließt verschiedene Arbeiten zur Modellierung von Textverstehen und Textverständlichkeit ein. Einige Arbeiten thematisieren die theoretische Unterscheidung von Sachstrukturen, Textstrukturen und Lernstrukturen (im Sinne des Verhaltens beim Lernen aus Texten). Diese Arbeiten sind überwiegend theoretischer Natur bzw. reflektieren den Stand der Forschung. Empirische Arbeiten befassen sich u. a. mit der Bedeutung des Interesses auf die Textverarbeitung oder die Untersuchung der psychometrischen Qualität zur Erfassung der Sprachfähigkeit von Schüler\*innen mit Migrationshintergrund.

**Thema 11 – Selbstkonzept** Das Selbstkonzept wird in Thema 11 doppelt behandelt: Einmal als Ergebnis von Unterricht und einmal als Prädiktor für die weitere Entwicklung der schulischen Leistung. Unter anderem wird die Rolle inkrementeller impliziter Fähigkeitstheorien auf die Entwicklung des Selbstkonzepts, aber auch auf das Wohlbefinden untersucht; oder etwa die Rolle impliziter Theorien bei der Wahrnehmung sozialer Eingebundenheit. Methodisch sind die Arbeiten zu diesem Thema primär empirischer Natur.

**Thema 12 – Förderung domänenspezifischen Wissens** Bei Thema 12 liegt der Fokus darauf, wie Lernende beim fach- bzw. domänenspezifischen Kompetenzerwerb unterstützt werden können. Erkenntnisziel der Forschung war dabei z. B. die Frage, inwiefern Handlungskompetenzen in einer bestimmten Domäne – beispielsweise der Buchführung – im Unterricht gefördert werden können, etwa durch die Verwendung von (ausgearbeiteten) Beispielen. Untersucht wird z. B. auch die Rolle instruktionaler Erklärungen bzw. Selbsterklärungen. Ein anderes Beispiel ist die Förderung des Leseverständnisses durch kognitive und meta-kognitive Strategien. Ziel dieses Themas ist damit die Untersuchung von Strategien zur Förderung unterschiedlicher domänenspezifischer Fähigkeiten und Fertigkeiten. Im Fokus liegt insbesondere die Identifikation von Schwierigkeiten der Lernenden bei der Nutzung solcher Strategien. Daraus sollen Erkenntnisse gewonnen werden, für welche Lernenden sich bestimmte Unterrichtsstrategien eignen bzw. wie die Strategien abhängig von den Lernenden angepasst werden müssen. Methodisch ist die Forschung im Thema 12 mehrheitlich experimentell angelegt.

**Thema 13 – Lernen im Tertiärbereich** Zusammenfassend befasst sich Thema 13 mit dem Lernen im Hochschulunterricht, der Bedeutung individueller und instruktionaler Faktoren für erfolgreiches Lehren und Lernen an Hochschulen. Zentral ist die Beschäftigung mit Faktoren, welche als Prädiktoren für den Studienerfolg gelten, z. B. allgemeine kognitive und motivational-affektive Voraussetzungen sowie weitere allgemeine Fähigkeiten und Fertigkeiten. Darunter fällt auch die Fähigkeit zum selbstregulierten Lernen. Ergänzend werden strukturelle bzw. organisatorische Aspekte der Hochschullehre berücksichtigt.

**Thema 14 – Wissenserwerb mit Multimedia** Thema 14 hebt auf das Lernen mit Multimedia, insbesondere den Wissenserwerb mit Hilfe von Multimedia ab. Insgesamt widmet sich Thema 14 dem Zusammenhang von multimedialer Informationsdarbietung und der Konstruktion mentaler Modelle beim Lernenden. Ziel ist die Etablierung eines Modells des Verstehens von Texten, Bildern und Diagrammen oder des computergestützten Lernens. Ebenfalls erforscht wird auch die Rolle multimedialer Lernumgebungen beim Problemlösen. Konkret wurde z. B. das Potenzial multimedialer Lernumgebungen für die Förderung der Kompetenz zum Lösen kategorialer, interpretativer und gestaltender Probleme untersucht. Dabei wurden verschiedene Varianten multimedialer Lernumgebungen experimentell verglichen – beispielsweise visuelle und audiovisuelle Varianten.

**Thema 15 – Nonverbale und verbale Kommunikation** Thema 15 thematisiert das kommunikative Verhalten angehender Lehrkräfte und Microteaching als Ansatz der Lehrkräftebildung. Ein erster Fokus lag auf dem nonverbalen kommunikativen Verhalten von (angehenden) Lehrenden. Ein Beispiel dafür ist das Training von Lehrverhalten in einer Fortbildung, beispielsweise die Theoriepräsentation. Für das Feedback zum nonverbalen kommunikativen Verhalten werden dazu verschiedene Arten von Beobachtungen erprobt. Dabei zeigt sich, dass expressives nonverbales kommunikatives Verhalten mit erhöhter Aufmerksamkeit der Zuhörer\*innen einhergeht. Andere Arbeiten analysieren verbales kommunikatives Verhalten: Dabei geht es z. B. um die Klarheit, Eindeutigkeit und Angemessenheit von Fragen der Lehrkraft. Die Ergebnisse des Trainings zeigen, dass danach vermehrt Fragen höherer Ordnung verwendet werden. Aufbauend auf der Studie wurden Trainingsmaterialien zur Förderung kommunikativen Fertigkeiten der Gesprächsführung entwickelt.

**Thema 16 – Selbstreflexion der Unterrichtsforschung** Thema 16 enthält Aufsätze, in denen eine Selbstreflexion der Unterrichtsforschung stattfindet. Es wird z. B. die Frage aufgeworfen, welche Rolle empirische Befunde für praktisches pädagogisches Handeln haben können. Andere Leitfragen entsprechender Aufsätze lauten: Wie soll Unterrichtsforschung methodisch betrieben werden? Ist es Aufgabe der Unterrichtsforschung, Wissen für die Praxis bereitzustellen? Insbesondere die Rolle der qualitativen Unterrichtsforschung wird diskutiert. Möglichkeiten und Erträge qualitativ ausgerichteter Unterrichtsforschung werden erörtert und dabei auch Verbindungen zur quantitativen Unterrichtsforschung diskutiert. Dazu gehören auch Arbeiten aus dem Bereich der videogestützten Unterrichtsforschung als Beispiel eines (neuen) methodischen Zugangs zur Unterrichtsforschung.



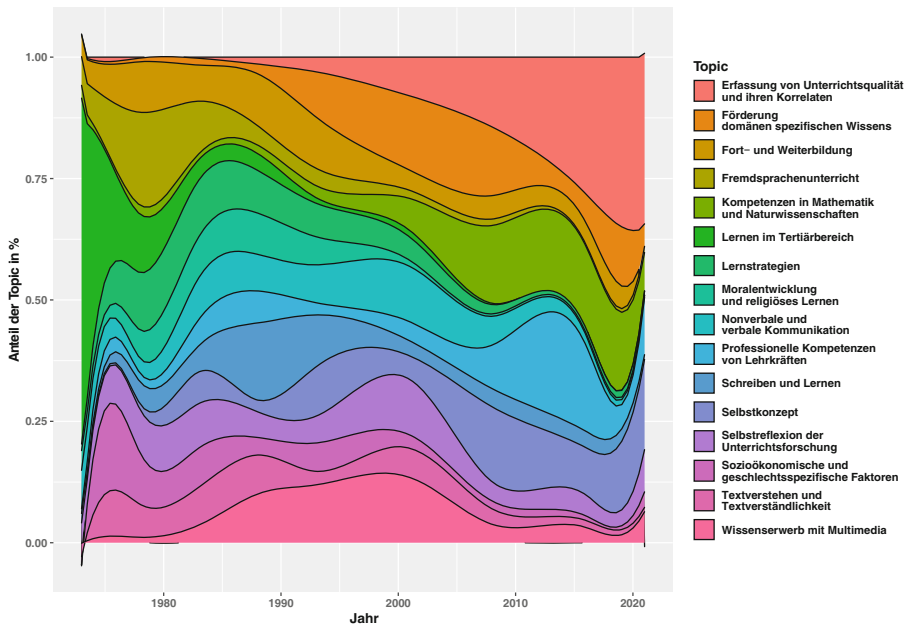


Abb. 4 Zeitliche Entwicklung der Themen als relative Häufigkeit über die Jahre von 1973 bis 2023

**Thema 17 – Moralentwicklung und religiöses Lernen** Im Fokus von Thema 17 stehen die moralische Entwicklung über die Lebensspanne (aufbauend auf den Arbeiten von Kohlberg) und die Bedeutung der Schule und Familie bzw. spezifischer pädagogischer Settings für die moralische Entwicklung. Das Thema berücksichtigt unterschiedliche Lernorte, welche zusammengenommen das lebenslange Lernen darstellen: Schule, Universität, Beruf, Familie und Massenkommunikationsmittel. Es zeigt sich zum Beispiel, dass eine moralische schulische Atmosphäre mit einem höheren moralischen Urteilsniveau assoziiert ist. Als ein Teilaspekt schließt diese Forschung auch religiöses Lernen ein. Beispielsweise zeigen die Untersuchungen analog zur Wirkung spezifischer moralisch orientierter Settings auf die moralische Entwicklung, dass ein religiöser Erziehungsstil mit höheren religiösen Urteilsstufen korreliert.

#### 4.2 Forschungsfrage 2: Wie haben sich diese Themen und damit die empirische Lehr-Lernforschung über die Zeit verändert?

Abb. 4 zeigt die Entwicklung der prozentualen Anteile der Topics pro Jahr über die 50 Jahre, in denen die *Unterrichtswissenschaft* existiert. Es lässt sich beobachten, dass die *Unterrichtswissenschaft* generell von einer großen thematischen Vielfalt geprägt ist. In den ersten Jahrgängen stand interessanterweise das Lernen im Tertiärbereich thematisch stark im Vordergrund. In der Zeit danach entfällt zu keinem Zeitpunkt mehr als ein Drittel des Inhalts auf ein einzelnes Thema. In dieser thematischen Vielfalt lassen sich jedoch klare Trendthemen identifizieren, die für eine

Zeitspanne einen größeren Anteil des Inhalts ausmachen und dann in ihrem Anteil am Inhalt wieder abnehmen oder sogar ganz verschwinden. So nimmt z. B. Topic 14 – Wissenserwerb mit Multimedia – parallel zur zunehmenden Verbreitung von Computern ab den 1980er-Jahren und durch die 1990er-Jahren hinweg stetig im Anteil zu und ab dem Jahr 2000 wieder ab. Für die letzten Jahre seit 2010 fällt auf, dass die thematische Vielfalt zurückgeht und sich Schwerpunkte auf die Topics 4 – Erfassung von Unterrichtsqualität und ihren Korrelaten, 7 – Professionelle Kompetenzen von Lehrkräften, 8 – Kompetenzen in Mathematik und Naturwissenschaften und 11 – Selbstkonzept ausprägen. Die übrigen Themen gehen im Anteil entsprechend zurück oder verschwinden sogar ganz. Auffallend ist, dass Topic 16, der sich mit der Selbstreflexion der Unterrichtsforschung beschäftigt, recht konstant bearbeitet wird, auch wenn hier ebenfalls ein Rückgang in den letzten Jahren zu verzeichnen ist. Seit 2020 zeichnet sich der Trend ab, dass Lernen mit Multimedia wieder wichtiger wird.

## 5 Diskussion und Reflexion

Der vorliegende Beitrag verfolgt im Wesentlichen zwei Ziele: 1) zentrale Themen der Zeitschrift *Unterrichtswissenschaft* und damit der deutschsprachigen Lehr-Lernforschung der letzten 50 Jahre und ihre Entwicklung zu identifizieren und 2) zur Bearbeitung dieser Frage eine Kombination aus quantitativem (computationalem) und qualitativem Zugang anzuwenden, die sogenannte *Computational Grounded Theory*. Im Folgenden diskutieren wir zunächst die inhaltlichen Befunde im Hinblick auf die eingangs dieses Beitrags formulierten Fragestellungen, bevor wir über den Einsatz der *Computational Grounded Theory* in diesem Zusammenhang reflektieren.

Die insgesamt 17 identifizierten Themen sind in ihrer Beschreibung im Großen und Ganzen schlüssig. Zwei Themen fallen dabei auf: Thema 1 ist ein Artefakt, bedingt durch die Struktur der vorliegenden Daten, bedingt durch die Struktur der vorliegenden Daten und wird daher weiter nicht mehr berücksichtigt. Thema 16 ist eher ein Metathema, das sich um die Frage der Erkenntnisse und der Fortentwicklung der empirischen Lehr-Lernforschung dreht. Die verbleibenden 15 Themen sind eher inhaltlicher Natur. Sie decken, über die Jahre, die verschiedenen möglichen Aspekte des Lehrens und Lernens in verschiedenen Institutionen gut ab – von der Ebene der Bedeutung individueller Merkmale (u. a. Themen 9 und 11) über die Lernumgebungen und deren Qualität (u. a. Themen 4 und 6) bis hin zur Ebene der Rahmenbedingungen, hier speziell der Lehrkräfteaus- und weiterbildung (u. a. Thema 3, 7 und 15). Die Ebene des Unterrichts zerfällt dabei in gewisser Weise in zwei größere Blöcke. Der erste Block umfasst eine Vielzahl unterschiedlicher Ansätze, das Lernen der Schülerinnen und Schüler im Unterricht (und in Teilen auch darüber hinaus) zu fördern. Hier werden häufig spezifische Aspekte des (unterrichtlichen) Lehrens und Lernens fokussiert, vor allem Instruktionsstrategien (u. a. Themen 2 und 10). Der zweite Block betrachtet Unterricht eher aus der Vogelperspektive und versucht Kriterien, Dimensionen oder auch Muster guten Unterrichts, sowohl in schulischen als auch anderen Settings, zu identifizieren (Thema 4). Auffallend bei der Betrachtung der Themen ist, dass (Kompetenzen in) Mathematik und Naturwis-

senschaften ihr eigenes Thema definieren (Thema 8). Es scheint, als würde dazu in den 1980er-Jahren, insbesondere aber seit dem Jahr 2000 eine ausreichende Menge spezifischer Beiträge publiziert. Dies erscheint mit Blick auf die Folgen der Befunde aus den großen internationalen Leistungsvergleichsstudien TIMSS und PISA und der darauffolgenden verstärkten Beschäftigung mit der Frage, über welche Kompetenzen Schüler\*innen in welchem Umfang verfügen, wenig überraschend und spiegelt den damaligen Trend vor allem der deutschen Unterrichtsforschung wider, die durch den „Schock“ des schlechten Abschneidens in der TIMSS und der ersten PISA-Studie angeregt wurde, sich intensiver mit Bedingungen des Kompetenzerwerbs in diesen Fächern zu befassen. Der Befund weist aber darauf hin, dass zu anderen Fächern nicht in gleichem Maße in der Lehr-Lernforschung publiziert wird. Eine alternative Erklärung könnte sein, dass entsprechende Beiträge nicht in ähnlicher Weise aufeinander Bezug nehmen (u. a. über die Verwendung derselben Theorien), wie dies in der Forschung zum Lehren und Lernen in Mathematik und Naturwissenschaften der Fall ist, und sie daher nicht in dieselbe Kategorie fallen.

Nicht zuletzt zeigt die Analyse Themen auf, die für einen gewissen Zeitraum, wenn auch nicht bestimmende, doch maßgebliche Themen der empirischen Unterrichts- bzw. Lehr-Lernforschung, wie z. B. die Moralentwicklung und -erziehung (Thema 17), die augenscheinlich die 1980er- bis 1990er-Jahre (mit)geprägt hat, oder auch die Multimedia-Forschung (Thema 14), die die 1990er- und 2000er-Jahre (mit)bestimmt hat. Betrachtet man diese Themen genauer, dann wird deutlich, dass die Themen der Zeitschrift *Unterrichtswissenschaft* auch durch die Forschungsinteressen einzelner Herausgeber\*innen geprägt war. Dies kann man z. B. für das Thema „Schreiben und Lernen“ (Forschungsthema von Gunther Eigler), multimediales Lernen (Heinz Mandl), Professionalisierung von Lehrpersonen (Jürgen Baumert) oder Unterrichtsqualität (Anna Praetorius) festhalten.

Mit Blick auf die Entwicklung der Themen fällt auf, dass sich in den letzten Jahren einige größere Themen herauskristallisiert haben – darunter die Unterrichtsqualität (Thema 4) sowie die Lehrkräfteaus- und weiterbildung (Themen 3, 7 und 15). Ähnlich wie bei Thema 8 (Mathematik und Naturwissenschaften) manifestieren sich hier die Folgen der großen Leistungsvergleichsstudien, darunter die intensive (videobasierte) Forschung zu Mustern guten Unterricht auf Basis und infolge der TIMSS-Videostudie(n) sowie in Folge dieser Arbeiten auch der zunehmende Fokus auf Lehrkräfte als die zentralen Akteure im Unterricht. Zuletzt ist das Forschungsthema der Unterrichtsqualität in seiner Bedeutung wieder stark gewachsen und seit einigen Jahren (wieder) das dominierende Thema. Dies mag konkret einerseits an einer Reihe von Sonderheften in den letzten Jahren liegen, ist andererseits aber mit großer Wahrscheinlichkeit in der Weiterentwicklung methodischer Zugänge von der Unterrichtsbeobachtung über Videoanalysen bis hin zu statistischen Verfahren, die die Analyse von Mustern überhaupt erst erlauben, begründet. In den beiden letzten Jahren werden wieder mehr Aufsätze zum Lernen mit Multimedia publiziert, möglicherweise findet hier das gesteigerte Interesse an der Digitalisierung für Lehren und Lernen ihren Niederschlag. Positiv fällt auch auf, dass über die Jahre kontinuierlich eine gewisse Selbstreflexion der unterrichtsbezogenen Lehr-Lernforschung stattgefunden hat (Thema 16). Dieses Thema hat über die gesamte Zeit hinweg – mit leichten Schwankungen, z. B. einem Anstieg um die Jahrtausendwen-

de – einen gewissen Stellenwert. Auffällig ist dabei, dass die Forschung, die stark auf die Entwicklung spezifischer Theorien des Lehrens und Lernens abhebt (d. h. wie Schülerinnen und Schüler lernen und welche konkreten Implikationen für den Unterricht bzw. unterrichtliche Strategien sich daraus ableiten, Thema 5), über die Zeit zunehmend abgenommen hat (siehe auch Themen 2, 6 und 10). Dies mag darin begründet liegen, dass sich mit dem Fokus weg von einzelnen Ergebnissen des unterrichtlichen Lehrens und Lernens (u. a. Wissen) hin zu einer mehr ganzheitlichen Zielsetzung (d. h. Kompetenz) auch der Schwerpunkt der Forschung zum unterrichtlichen Lehren und Lernen von einem eher engen Fokus auf das Lernen bestimmter Fähigkeiten und entsprechender Instruktionsstrategien auf einen eher breiten Zugang im Sinne der Betrachtung der Qualität von Unterricht als Ganzes verschoben hat. Es mag aber auch einfach ein Zeichen für einen zunehmenden Mangel an theoretischer Fundierung bzw. einen weniger starken Theoriebezug sein.

Die beschriebenen Erkenntnisse sind in dieser Form also das Resultat der Anwendung der neuen Methode der *Computational Grounded Theory*. Die vorgestellten Themen und ihre zeitliche Entwicklung wären ohne die verschränkte Verwendung von Verfahren des maschinellen Lernens (oder künstlicher Intelligenz im Allgemeinen) und menschlicher Analyse nicht möglich gewesen und zeigen so, dass *Computational Grounded Theory* eine Bereicherung im Methodenrepertoire darstellt. Ob die Bereicherung, wie von Nelson (2020) dargelegt, in der verbesserten Objektivität und Reproduzierbarkeit von qualitativen Analysen liegt, bleibt allerdings zu bezweifeln. Vielmehr scheint eine Stärke der Methode zu sein, die Stärken von Mensch und Maschine so zu kombinieren, dass sie sich bestmöglich ergänzen. Dies wird klar, wenn man sich den Umfang der Analysen in Erinnerung ruft: Es ist nur sehr eingeschränkt möglich, dass ein Forschungsteam oder gar ein Mensch alle Abstracts aller Jahrgänge der Unterrichtswissenschaft liest und hierbei über *alle* Jahrgänge hinweg Muster über verschiedene Textstellen hinweg identifiziert. Der schiere Umfang des Textes und die Begrenztheit des menschlichen (Arbeits)gedächtnisses bzw. die Begrenztheit der Mustererkennung durch distinkte Analysekategorien schließen dies aus – gerade wenn man an subtile Muster denkt, die weit verteilt in den Daten vorhanden sind. Diese Muster kann ein Computer mit seinen Speicherkapazitäten und Analysemöglichkeiten durchaus erkennen. Allerdings ist dem Computer eine inhaltliche Deutung der gefundenen Muster nicht möglich. Hier kann wiederum der Mensch, basierend auf einer vom Computer getroffenen Vorauswahl, seine Fähigkeiten zur Interpretation zum Tragen bringen. In diesem Sinne ergänzen sich Mensch und Computer im epistemischen Prozess (vgl. Baumer et al. 2017; Kubsch et al. 2022). Wie wichtig diese Ergänzung ist, zeigt auch Thema 1, welches ein Artefakt darstellt. Hier werden bedingt durch die Struktur der für diese Arbeit vorliegenden Rohdaten hauptsächlich Literaturverzeichnisse abgebildet. Dass dies erkannt wurde, ist der menschlichen Interpretation der Texte zu verdanken – aus Sicht des Computers ist Thema 1 so gut wie jedes andere der gefundenen Themen. Gleichzeitig ist die Identifikation von Thema 1 als Artefakt ein Beleg, dass *Computational Grounded Theory* als Prozess so angelegt ist, dass ein Schutz vor solchen Artefakten und daraus möglichen Fehlschlüssen gegeben ist.

Ist dieser Prozess durch den Einsatz eines Computers dabei objektiver und reproduzierbarer geworden? Während die Reproduzierbarkeit durch das Vorhalten von

Analyseskripten zumindest prinzipiell gewährleistet werden kann, ist die größere Objektivität aufgrund der vielen Optionen in den Computerverfahren, die von den Forscher\*innen gewählt werden können, nicht unbedingt gegeben. Wir mussten diverse Entscheidungen in der Datenaufbereitung und -analyse treffen, welche gut begründet auch anders hätten getroffen werden können (vgl. Baumer et al. 2017; Simmons et al. 2011). Zum Beispiel hätte man weitere stopwords hinzufügen oder andere Hyperparameter für die Modelle wählen können. Prinzipiell könnten sogenannte Multiverse Analysen (Steege et al. 2016), welche alle Wahlmöglichkeiten durchführen und am Ende die Ergebnisse mitteln, dem Problem entgegenwirken, allerdings kommt man bei dem Umfang der hier verwendeten Daten auch schnell an die Grenzen des Realisierbaren. Es folgt, dass ein höherer Grad von Objektivität durch *Computational Grounded Theory* nur insofern gegeben sein kann, wie der Computereinsatz ansonsten vorhandene Freiheitsgerade einschränkt oder Entscheidungen an objektivierbare Maßstäbe, wie zum Beispiel Fitindizes, knüpft.

Zusammenfassend hat die Verknüpfung quantitativer (computationaler) und qualitativer Verfahren im Rahmen der *Computational Grounded Theory* ein hohes Potenzial für die Weiterentwicklung der empirischen Lehr-Lernforschung. In ihr treffen verschiedene epistemische Zugänge aufeinander – die interpretative Tradition der qualitativen Analyse trifft auf den Zugang der quantitativen (computationalen) Verfahren (vgl. auch Baumer et al. 2017). Die potenziell daraus resultierenden epistemischen Spannungen können produktiv für den Erkenntnisgewinnungsprozess werden, wenn sie zu einer kritischen Aufmerksamkeit und Reflektivität der Forschenden führen. Im vorliegenden Beitrag konnten mit Hilfe der *Computational Grounded Theory* vorurteilsfrei aus den Texten heraus 17 zentrale Themen der empirischen (unterrichtsbezogenen) Lehr-Lernforschung identifiziert werden. Manche dieser Themen wären sicher auch in einem traditionellen Zugang identifiziert worden, andere möglicherweise nicht, weil sie nicht mehr aktuell sind (z. B. Thema 17: Moralentwicklung und religiöses Lernen) oder in anderen Themen aufgegangen wären (z. B. Thema 15: nonverbale und verbale Kommunikation). Natürlich sind die hier vorgestellten Beschreibungen nach wie vor durch eine gewisse Subjektivität geprägt. Die identifizierten Themen, ihr Zustandekommen und die ihnen zugeordneten Texte liegen jedoch vor und sind einer Prüfung unmittelbar zugänglich. Nicht zuletzt bieten insbesondere die quantitativen Analysen der beschriebenen Themen interessante Einblicke in die Entwicklung des Feldes. Für regelmäßige Leser\*innen dürfte die Entwicklung der beiden großen Themen Unterrichtsqualität und Lehrkräfteaus- und -weiterbildung zu eben solchen zwar keine Überraschung sein, überraschend ist dagegen die Erkenntnis, dass stärker fokussierte, stark theoriebasierte Forschung zum Lernen bzw. zu Lernprozessen und zur Steuerung von Lernprozessen über die Zeit stark abgenommen hat. Dies wird kritisch zu reflektieren sein.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Interessenkonflikt** M. Kubsch, K. Neumann, M. Rochnia und C. Gräsel geben an, dass kein Interessenkonflikt besteht.

**Open Access** Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ord-

nungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

## Literatur

- Bail, C.A. (2014). The cultural environment: Measuring culture with big data. *Theory and Society*, 43(3–4), 465–482. <https://doi.org/10.1007/s11186-014-9216-5>.
- Baumer, E.P.S., Mimno, D., Guha, S., Quan, E., & Gay, G.K. (2017). Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence? *Journal of the Association for Information Science and Technology*, 68(6), 1397–1410. <https://doi.org/10.1002/asi.23786>.
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30), 774. <https://doi.org/10.21105/joss.00774>.
- Biernacki, R. (2012). *Reinventing evidence in social inquiry: Decoding facts and variables* (1st ed.). Palgrave Macmillan.
- Blei, D.M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77. <https://doi.org/10.1145/2133806.2133826>.
- Blei, D.M., & Lafferty, J.D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, 1(1), 17–35. <https://doi.org/10.1214/07-AOAS114>.
- Christianson, N.H., Sizemore Blevins, A., & Bassett, D.S. (2020). Architecture and evolution of semantic networks in mathematics texts. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 476(2239), 20190741. <https://doi.org/10.1098/rspa.2019.0741>.
- Glaser, B.G., & Strauss, A.L. (2017). *The discovery of grounded theory: Strategies for qualitative research* (1. Aufl.). Routledge. <https://doi.org/10.4324/9780203793206>.
- Gräsel, C., & Gniewosz, B. (2011). Überblick Lehr-Lernforschung. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung*. VS. [https://doi.org/10.1007/978-3-531-93021-3\\_1](https://doi.org/10.1007/978-3-531-93021-3_1).
- Jank, W., & Meyer, H. (2002). *Didaktische Modelle* (5. Aufl.). Berlin: Cornelsen-Scriptor.
- Kiel, E. (Hrsg.). (2018). *Unterricht sehen, analysieren, gestalten*. UTB.
- Kleickmann, T., & Hardy, I. (2019). Vernetzung professionellen Wissens angehender Lehrkräfte im Lehramtsstudium. *Unterrichtswissenschaft*, 47(1), 1–6.
- Kubsch, M., Krist, C., & Rosenberg, J.M. (2022). Distributing epistemic functions and tasks—A framework for augmenting human analytic power with machine learning in science education research. *Journal of Research in Science Teaching*. <https://doi.org/10.1002/tea.21803>.
- Kuckartz, U., & McWhertor, A. (2014). *Qualitative text analysis: A guide to methods, practice & using software*. SAGE.
- Landauer, T.K. (2014). *Handbook of latent semantic analysis*. Routledge.
- Mayring, P. (2014). *Qualitative content analysis*. Beltz.
- Mimno, D., Wallach, H.M., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (S. 262–272).
- Nelson, L.K. (2020). Computational grounded theory: a methodological framework. *Sociological Methods & Research*, 49(1), 3–42. <https://doi.org/10.1177/0049124117729703>.
- Odden, T.O.B., Marin, A., & Caballero, M.D. (2020). Thematic analysis of 18 years of physics education research conference proceedings using natural language processing. *Physical Review Physics Education Research*, 16(1), 10142. <https://doi.org/10.1103/PhysRevPhysEduRes.16.010142>.
- Odden, T.O.B., Marin, A., & Rudolph, J.L. (2021). How has *Science Education* changed over the last 100 years? An analysis using natural language processing. *Science Education*, 105(4), 653–680. <https://doi.org/10.1002/sce.21623>.

- Praetorius, A. K., & Nehring, A. (2020). Unterrichtsqualität zwischen Generik und Fachspezifik: Einführung in den Thementeil. *Unterrichtswissenschaft*, 48(3), 297–301.
- Rittelmeyer, C., Parmentier, M., & Klafki, W. (2001). *Einführung in die pädagogische Hermeneutik*. WBG.
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). *stm*: An R package for structural topic models. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v091.i02>.
- Rosenberg, J. M., & Krist, C. (2020). Combining machine learning and qualitative methods to elaborate students' ideas about the generality of their model-based explanations. *Journal of Science Education and Technology*. <https://doi.org/10.1007/s10956-020-09862-4>.
- Rost, M., Barkhuus, L., Cramer, H., & Brown, B. (2013). Representation and communication: Challenges in interpreting large social media datasets. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (S. 357–362).
- Schröder, B. (2012). Vergleichende historische Religionspädagogik – methodologische Überlegungen. *Zeitschrift für Pädagogik und Theologie*, 68(2), 238–251.
- Shaw, E., Phillips, R., Kao, A., & Torres, R. (2016). Using topic modeling to assess middle school science discourse based on next generation science standards. In *Proceedings of The 2017 International Conference on Advanced Technologies Enhancing Education (ICAT2E 2017)*. *The 2017 International Conference on Advanced Technologies Enhancing Education (ICAT2E 2017)*, Qingdao, China. <https://doi.org/10.2991/icat2e-17.2016.11>.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>.
- Steege, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, 11(5), 702–712. <https://doi.org/10.1177/1745691616658637>.
- Zlatkin-Troitschanskaia, O., Pant, H. A., & Blömeke, S. (2016). Herausforderungen und Erfolge in der Modellierung und Messung akademischer Kompetenzen. *Unterrichtswissenschaft*, 44(1), 2–6.