

Fachartikel aus:

Berenike Herrmann / Maria Kraxenberger (Hg.): Weder Fail noch Lobgesang. Nichteindeutige Wertung von Literatur im digitalen Raum (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 6). 2025. DOI: [10.17175/sb006](https://doi.org/10.17175/sb006)

Titel:

»Einfach klasse!« Eine Mixed-Methods-Analyse zur Eindeutigkeit literarischer Wertungspraktiken auf Amazon und BücherTreff.de

Autor*in:

Anna Moskvina

Kontakt: moskvi@uni-hildesheim.de

Institution: Universität Hildesheim / Ella Lab – Gesellschaft für künstliche Intelligenz

GND: [1324505389](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63337-2101) ORCID: [0009-0006-3337-2101](https://orcid.org/0009-0006-3337-2101)

Contribution (CRediT): [Data curation](#) | [Formal analysis](#)

Autor*in:

Kristina Petzold

Kontakt: kristina.petzold@uni-bielefeld.de

Institution: Universität Bielefeld / SFB 1288 "Praktiken des Vergleichens", Teilprojekt D05 „Vergleichendes Lesen“

GND: [1197555811](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63337-2101) ORCID: [0009-0006-3337-2101](https://orcid.org/0009-0006-3337-2101)

Contribution (CRediT): [Conceptualization](#) | [Writing – original draft](#) | [Formal analysis](#)

DOI des Beitrags:

[10.17175/sb006_004](https://doi.org/10.17175/sb006_004)

Nachweis im OPAC der Herzog August Bibliothek:

[1928799159](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63337-2101)

Erstveröffentlichung:

20.11.2025

Lizenz:

Sofern nicht anders angegeben 

Letzte Überprüfung aller Verweise:

20.06.2025

Format:

PDF ohne Paginierung, Lesefassung

GND-Verschlagwortung:

[Methodenmix](#) | [Rezension](#) | [Literaturwissenschaft](#) | [Eindeutigkeit](#) | [Werturteil](#)

Empfohlene Zitierweise:

Anna Moskvina / Kristina Petzold: »Einfach klasse!« Eine Mixed-Methods-Analyse zur Eindeutigkeit literarischer Wertungspraktiken auf Amazon und BücherTreff.de. In: Berenike Herrmann / Maria Kraxenberger (Hg.): Weder Fail noch Lobgesang. Nichteindeutige Wertung von Literatur im digitalen Raum (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 6). Wolfenbüttel 2025. 20.11.2025. HTML / XML / PDF. DOI: [10.17175/sb006_004](https://doi.org/10.17175/sb006_004)

Anna Moskvina / Kristina Petzold

»Einfach klasse!« Eine Mixed-Methods-Analyse zur Eindeutigkeit literarischer Wertungspraktiken auf Amazon und BücherTreff.de

Abstract

Im Beitrag werden Online-Rezensionen zu Büchern verschiedener literarischer und nichtliterarischer Genres mit dem Ziel analysiert, Eindeutigkeiten und Nichteindeutigkeiten von digitalen Wertungspraktiken genauer zu beschreiben. Es wird danach gefragt, inwiefern Eindeutigkeit als mögliches Unterscheidungskriterium zwischen Bewertungspraktiken auf verschiedenen Plattformen und zwischen numerischen und textuellen Wertungspraktiken fungieren kann. Online-Rezensionen von Amazon (deutschsprachig) und BücherTreff.de werden dabei mithilfe dreier verschiedener Analyseverfahren (*Codierungs-Statistik*, *Sentiment-Analyse* und *qualitativer Inhaltsanalyse*) untersucht.

This article analyses online reviews of books of different literary and non-literary genres with a focus on the clarities and ambiguities of digital evaluation practices. The authors ask to what extent clarity can function as a possible distinguishing criterion between evaluation practices on different platforms and between numerical and textual practices of evaluation. Online reviews from Amazon (German corpus) and BücherTreff.de are examined using three different analytical methods (*coding statistics*, *sentiment analysis* and *qualitative content analysis*).

1. Lob und Torheit – zur Einleitung

Ein gängiges Vorurteil gegenüber digitalen Buchrezensionen von Lai*innen besteht in der Annahme, dass diese vorwiegend positiv und undifferenziert – ein bloßer ›Lobgesang‹ – seien.¹ Vieles an dieser Behauptung ist problematisch: Erstens deutet die bisherige Forschung darauf hin, dass es nicht sinnvoll ist, digitale Rezensionsformen, die in sehr unterschiedlichen medialen und sozialen Kontexten auftreten, allzu stark zu verallgemeinern (siehe hierzu den Forschungsüberblick in Kapitel 2). Im Gegenteil scheint eine Differenzierung mindestens nach Plattform und rezensiertem Artefakt sinnvoll, vermutlich aber auch nach spezifischen *communities of practice*.

[1]

Zweitens stellt sich die Frage, inwiefern man bei Buchrezensent*innen auf Online-Plattformen, in sozialen Netzwerken und auf *Social-Reading*-Portalen von Lai*innen sprechen kann, wenn zum einen die Identität und damit auch die Profession gar nicht immer sichtbar sind (z. B. auf Amazon) und wenn zudem unklar bleibt, worauf die ›Laienhaftigkeit‹ denn bezogen wird.² Dass damit häufig auf mangelnde institutionelle Qualifikations- und Legitimationsnachweise sowie normative Textsortenkriterien einer spezifischen Rezensionsform (der Feuilleton-Kritik) angespielt wird, legen Fälle nahe, in denen eine sehr hohe Professionalität von Online-Rezensent*innen in Bezug auf den besprochenen Gegenstand (z. B. Genre-Literatur) oder hinsichtlich ihrer Medienkompetenz (z. B. Videotechnik und Plattformfunktionen betreffend) bei der Einordnung solcher Personen als ›Lai*innen‹ keine Rolle spielen.

[2]

Drittens gibt es zwar inzwischen empirische Belege dafür, dass beispielsweise Buchrezensionen auf Amazon tatsächlich in der Regel positiv ausfallen, quantitative Studien zur Stilistik solcher Rezensionen, zu ihren inhaltlichen Schwerpunkten und Bewertungspraktiken sind – insbesondere im deutschsprachigen Raum – aber weiterhin rar.³ Es kann also bisher nicht mit Sicherheit festgehalten werden, dass solche Rezensionen

[3]

¹ Zum Diskurs über Buchblogger*innen sowie insbesondere den Abwertungsstrategien in Bezug auf diese, siehe Petzold 2025. Überdies scheint sich im literarischen Feld die Verknüpfung von Trivialität und Eindeutigkeit auch über moderne / postmoderne Definitionen des Literarischen bzw. des künstlerisch Anspruchsvollen anzubieten, das als idealerweise komplex und mehrdeutig aufgefasst wird (vgl. Schmidt 1991; Luhmann 1997; Baßler / Drügh 2021).

² Vgl. Giacomuzzi 2021.

³ Als positive Beispiele sind hier die Amazon-Studien von Bachmann-Stein 2015 und Mehling et al. 2018 zu nennen.

besonders undifferenziert sind oder dass beispielsweise numerisch maximal positive 5-Sterne-Rezensionen nicht auch Kritikpunkte an ihrem Rezensionsobjekt textuell äußern. Anders formuliert: Ob die hohen Anteile positiver Wertungen auf Online-Plattformen tatsächlich auf *eindeutig positiven* Wertungen basieren, ist bisher nicht untersucht worden. Dabei scheint gerade die Frage nach der Eindeutigkeit oder Nichteindeutigkeit von Online-Rezensionen viele Anschlussmöglichkeiten für die oben diskutierten Zuschreibungen zu ermöglichen.

Der vorliegende Beitrag widmet sich daher der Frage, ob Online-Rezensionen auf Amazon und auf der deutschsprachigen Leseplattform BücherTreff.de eindeutig oder nichteindeutig werten. Die Datenbasis bilden umfangreiche Textkorpora der beiden Plattformen, die mithilfe manueller und maschineller Analyseverfahren untersucht wurden. Dabei wurde in einem ersten Schritt ein händisches Codierungsverfahren durchgeführt, welches der Identifikation von eindeutig positiven, eindeutig negativen, nichteindeutigen und neutralen Bewertungen innerhalb von Online-Rezensionen diene. Die statistische Auswertung dieser Codierungen wurde im Anschluss mit *Polarity-Wortlisten*⁴ abgeglichen, was sowohl weitere Aussagen über die Urteilstendenz von Online-Rezensent*innen auf den untersuchten Plattformen zuließ als auch über das Verhältnis der beiden quantitativen Verfahren zueinander und damit deren jeweilige Vor- und Nachteile. In einem nächsten Schritt wurden die mithilfe statistischer Exploration ausgewählten Rezensionen qualitativ analysiert, um Erkenntnisse darüber zu erlangen, welche textuellen Aspekte zur Nichteindeutigkeit oder Eindeutigkeit von Online-Rezensionen beitragen können. Abschließend wurden die Ergebnisse der drei verschiedenen Verfahren kritisch reflektiert sowie weiterführende Fragestellungen in Bezug auf die Eindeutigkeit von Wertungen in Online-Rezensionen diskutiert.

[4]

2. Vorüberlegungen zur Eindeutigkeit in Online-Rezensionen

In den letzten Jahren waren Online-Rezensionen zunehmend Gegenstand der Forschung.⁵ Der Einfluss von Rezensionstexten auf Kaufentscheidungen wurde untersucht⁶ sowie deren Bedeutung als sogenannter *user generated content*.⁷ Insbesondere bei Buchrezensionen wurde der Fokus auf stilistische und inhaltliche Aspekte gelegt, um Online-Rezensionen als eine Form der Laien-Literaturkritik zu beschreiben.⁸ Darüber hinaus wird die wachsende Menge an verfügbaren Daten für eine allgemeinere, empirische Rezeptionsforschung⁹ sowie für methodologische Arbeiten genutzt.¹⁰ Aus diesen und anderen Studien lässt sich u. a. ableiten, dass sich Online-Rezensionen formal und inhaltlich nach Plattformen unterscheiden,¹¹ dass aber auch das Rezensionsobjekt bzw. dessen Genre¹² sowie das Selbstverständnis und die Erfahrung der Rezensent*in einen Einfluss auf Inhalt und Form von Online-Rezensionen haben.¹³ Insbesondere zu Amazon-Rezensionen liegen für den deutschsprachigen Raum Studien vor, die diese als vorwiegend positiv und nicht an Kriterien professioneller Kritik orientiert, sondern als erlebnisbezogen beschreiben.¹⁴

[5]

⁴ SentiWS, Remus 2010.

⁵ Ein systematischer Überblick über die aktuelle DH-Forschung zu verschiedenen Phänomenen des Social Reading findet sich bei Rebora et al. 2021.

⁶ Vgl. Chevalier / Mayzlin 2006; Kerstan 2006; Sutton / Paulfeuerborn 2017.

⁷ Vgl. u. a. Bruns 2008; Kirchmeier 2018.

⁸ Vgl. u. a. für deutschsprachige Rezensionen Anz 2010; Kellermann / Mehling 2017; Neuhaus 2015. In der internationalen Forschung wurde die Thematik u. a. adressiert von Allington 2016; McDonald 2007; Johnson 2016.

⁹ Vgl. u. a. Driscoll / Rehberg Sedo 2018; Lendvai et al. 2019. Im Sammelband von Moser / Dürr (Hg.) 2021 finden sich ebenfalls mehrere Beiträge dazu.

¹⁰ Vgl. u. a. Rebora et al. 2018.

¹¹ Vgl. Dimitrov et al. 2015; Koolen et al. 2020; Moskvina / Petzold [im Erscheinen].

¹² Vgl. Koolen et al. 2020; Boot / Koolen 2020.

¹³ Vgl. Koolen et al. 2020; Reinwand-Weiss / Roßkopf 2021, S. 84–93.

¹⁴ Vgl. Bachmann-Stein 2015; Kellermann / Mehling 2017; Porombka 2011; Stein 2015.

Vor dem Hintergrund dieser Studien werden Online-Rezensionen nachfolgend als kulturelle Praktiken verstanden,¹⁵ die in Abhängigkeit von verschiedenen Faktoren, wie den *affordances* und *constraints* ihres medialen Äußerungsortes, gruppenbezogenen Konventionen, dem Rezensionsobjekt, stabilen und situativen Dispositionen der individuellen Verfasser*in und deren Selbstverständnis, je unterschiedliche Funktionen erfüllen und unterschiedliche Inhalte und Formen aufweisen können. Unsere Hypothese zur Eindeutigkeit von Bewertungen in Online-Rezensionen kann dabei nicht alle diese Faktoren umfassen. Sie besteht daher in der Vermutung, dass die Eindeutigkeit der Bewertung abhängig ist von der Plattform, auf der die Rezension veröffentlicht wird und sich insofern die Rezensionen auf Amazon und BücherTreff.de hinsichtlich ihrer Eindeutigkeit unterscheiden. Besonders im Rahmen der qualitativen Auseinandersetzung mit exemplarischen Rezensionstexten im dritten Teil dieser Studie werden zudem auch andere Einflussfaktoren datengeleitet in den Blick genommen.

[6]

Wertungshandlungen bzw. wertende Sprechhandlungen stellen innerhalb der Rezensionen auf beiden Plattformen eine – wenn nicht *die* – zentrale inhaltliche Komponente dar.¹⁶ Als Wertung wird mit Renate von Heydebrand und Simone Winko eine implizite oder explizite Sprachhandlung verstanden, durch die einem Referenzobjekt ein bestimmter attributiver Wert (basierend auf einem axiologischen Wert) zugewiesen wird.¹⁷ Dass dabei auch Nichteindeutigkeiten entstehen können, ist insofern zu erwarten, als Unbestimmtheiten und Unterspezifikationen in natürlichen Sprachen nicht die Ausnahme, sondern eher den Normalfall bilden.¹⁸ Da im Rahmen des Beitrages eine systematische Definition der unterschiedlichen Formen von Nichteindeutigkeiten – wie Mehrdeutigkeit / Ambiguität, Vagheit und Kontextabhängigkeit – zu weit führen würde, soll der umgekehrte Weg gewählt werden und die folgende Arbeitsdefinition eindeutiger Wertungen als Grundlage dienen: Eine eindeutige Wertung liegt dann vor, wenn im vorliegenden sprachlich-pragmatischen Kontext einem Referenzobjekt ein attributiver Wert eindeutig zugeordnet wird und keine anderen Lesarten oder die Wertung in Frage stellende Interpretationsspielräume identifiziert werden können.

[7]

Mit dieser Arbeitsdefinition kann erstens der (mediale) Kontext als Faktor der ›Vereindeutigung‹ berücksichtigt werden. Gerade auf Online-Plattformen enthalten Seitenstruktur und -layout häufig wichtige Informationen (wie Buchtitel, Genre oder Autor*in), die dann im Rezensionstext nicht immer explizit erwähnt werden. Dennoch ist in solchen Fällen entsprechend der Definition von Eindeutigkeit auszugehen. Zweitens bleibt durch die wirkungsbezogene Operationalisierung anhand der Frage, wo konfligierende Lesarten bzw. widersprüchliche Interpretationen denkbar sind, die Liste der Einflussfaktoren eine offene. Welche Komponenten die sprachliche Etablierung einer eindeutigen Wertung (bewusst oder unbewusst sowie implizit oder explizit) kompromittieren können, wird somit erst durch die Analyse beantwortet. Auch, welche Formen der Nichteindeutigkeit in Online-Rezensionen dominieren bzw. bevorzugt eingesetzt werden, ist Teil dieser Fragestellung.

[8]

Um die Hypothese zu prüfen, dass Online-Rezensionen auf unterschiedlichen Plattformen sich hinsichtlich ihrer Eindeutigkeit unterscheiden, werden Rezensionen der beiden Plattformen Amazon und BücherTreff.de miteinander verglichen. Bei beiden handelt es sich um öffentliche und potenziell anonyme Online-Plattformen, auf denen Nutzer*innen neben anderen Textsorten auch Buchrezensionen verfassen können, welche dort langfristig abrufbar sind.¹⁹ Die Plattformstruktur orientiert sich in beiden Fällen primär an den rezensierten Büchern (bzw. anderen Produkten) und weniger an einzelnen Nutzer*innenprofilen, wie es beim klassischen *Social Network* oder auf vielen Social-Reading-Plattformen mit den entsprechenden Formen

[9]

¹⁵ Vgl. zur handlungsorientierten Auffassung von Rezensionen bzw. zur praxeologischen Rezeptionsforschung von Heydebrand / Winko 1996; Knipp 2017.

¹⁶ Dafür sprechen sowohl normative Textsortendefinitionen wie z. B. bei Anz 2004, S. 228, ebenso wie die weiter oben bereits zitierte deskriptive Forschung zu Online-Rezensionen.

¹⁷ Vgl. von Heydebrand / Winko 1996, S. 47–48.

¹⁸ Vgl. Römer / Matzke 2005, S. 137.

¹⁹ Die Beschreibungskategorien entstammen verschiedenen Taxonomien des Social Reading u. a. bei Stein 2012; Kutzner et al. 2019; Perschak 2021.

des *Social Cataloguing* der Fall ist.²⁰ Die Verbindungen zwischen einzelnen Rezensionstexten sowie der Interaktionsgrad zwischen den Rezensent*innen wird als eher gering eingeschätzt.²¹ Ein wichtiger Unterschied zwischen den Plattformen besteht darin, dass Amazon einen klar verkaufsorientierten Ansatz verfolgt, während es sich bei BücherTreff.de um eine »verlagsunabhängige[] Buch-Community«²² handelt. Das heißt, anders als im Falle von Social-Reading-Plattformen wie [LovelyBooks](#) sind die Betreiber nicht Verlage, sondern Privatpersonen. Die Finanzierung erfolgt über Bannerwerbung, Verlagskooperationen in Form von gesponserten Leserunden und *Affiliate-Links* u. a. zu Amazon, Thalia und Medimops.²³

Eine Ergänzung der Arbeitshypothese bildet daher die Vermutung, dass die spezifische Nutzungsumgebung von Plattformen, die stärker auf Quantifizierung (durch Sternevergabe) und unmittelbare Kaufentscheidungen ausgerichtet sind (wie Amazon), eindeutige Rezensionen hervorbringt. Umgekehrt ließe sich für nichtkommerzielle Plattformen wie BücherTreff.de, die eher austausch- und diskussionsorientiert sind und auf denen quantifizierende Verfahren nur optional angeboten werden, prognostizieren, dass diese mehr Nichteindeutigkeiten in ihren Buchbewertungen aufweisen. [10]

3. Datenbasis und Vorgehen

Die Grundlage der statistischen Auswertung zur Eindeutigkeit von Bewertungen in Online-Rezensionen bildet das Codiersystem des Forschungsprojektes Rez@Kultur.²⁴ Ein erster wichtiger Befund in Bezug auf die Eindeutigkeit von Wertungen in Online-Rezensionen besteht hier bereits darin, dass das induktiv erstellte Kategoriensystem eine Kategorie der nichteindeutigen Wertung überhaupt enthält. In der Logik des Kategoriensystem handelt sich dabei um ein Attribut innerhalb der Unterkategorie »Werten«, welche wiederum eine der zentralen Sprachhandlungen in Online-Rezensionen darstellt.²⁵ Die Datensets, die zur Bildung der Kategorien herangezogen wurden, umfassten Rezensionen zu Werken der Bildenden Kunst sowie Buchrezensionen und kurze empfehlende Texte auf Online-Handelsplattformen ebenso wie lange essayistische Formen (u. a. Buchblogs).²⁶ In dieser heterogenen Sammlung erwies sich die Markierung von nichteindeutigen Wertungen als ein relevanter inhaltlicher Aspekt von rezensiven Texten. Hieraus lässt sich ein erster Hinweis darauf ablesen, dass Nichteindeutigkeit im Urteil über ein künstlerisches Artefakt in sehr unterschiedlichen digitalen Rezensionsformen vorkommt und dass es sich dabei tendenziell nicht um das distinkte Merkmal von beispielsweise Lai*innenrezensionen handelt oder spezifisch ist für die Wertung von Literatur. [11]

²⁰ Vgl. Perschak 2021, S. 111.

²¹ Vgl. Perschak 2021, S. 111. Die These basiert auf Perschaks vergleichender Untersuchung zu Leserunden und Bucherdiskussionen, wobei sie letzteren eine geringere Interaktionsrate zuschreibt. Auf BücherTreff.de werden neben Rezensionen auch stärker interaktionale Formate wie Leserunden angeboten, die allerdings in der vorliegenden Studie keine Rolle spielen.

²² [BücherTreff.de](#) 2024.

²³ Vgl. Börsenblatt 2016.

²⁴ Dieses Codiersystem wurde durch die iterative Analyse von Online-Rezensionen entsprechend der Grounded Theory Methodology (Glaser / Strauss 2010 bzw. Strauss / Corbin 1996) datengetrieben erarbeitet. In mehreren Erhebungen wurden dazu Online-Rezensionen von verschiedenen Plattformen qualitativ analysiert und im Team von ca. sechs Forscher*innen verschiedener Disziplinen diskutiert und evaluiert (vgl. Kutzner et al. 2018). Ziel war es, typische Referenzierungen und Sprachhandlungen solcher rezensiver Texte mithilfe eines gemeinsamen Vokabulars zu identifizieren. Das Ergebnis dieses iterativen Vorgehens bildet ein Kategoriensystem bzw. Tagset, das zur händischen Codierung von rezensiven Texten über Literatur und bildende Kunst verwendet werden kann. Das vollständige Kategoriensystem ist im Abschlussband des Rez@Kultur-Projektes von Graf et al. (Hg.) 2021 dokumentiert. Da alle Kategorien und die annotierten Teile eines Rezensionstextes das Ergebnis einer Diskussion zwischen den Forschern waren, wurde keine Übereinstimmung zwischen den Annotatoren berechnet.

²⁵ Aufgrund der zugrundeliegenden Textsortenvielfalt ist das keineswegs eine triviale Feststellung.

²⁶ Dieses heterogene Korpus wurde lediglich in einem ersten Schritt, zur Erstellung des Kategoriensystems, genutzt, während der zweite Schritt, die quantitative Anwendung der Kategorien, sich auf Rezensionen von BücherTreff.de und Amazon konzentrierte.

Werten (10.3)		
Positiv/Zustimmung (10.3.1)	Wird ein Aspekt (z. B. aus Ebene 1) positiv gewertet?	Korpusschnitt: »Maggie McGinnis Schreibstil fand ich toll.« (BücherTreff.de, Rez 2171087)
Negativ/Widerspruch (10.3.2)	Wird ein Aspekt (z. B. aus Ebene 1) negativ gewertet?	Korpusschnitt: »Leider ein gähnend langweiliger Möchtegern-Thriller mit unsympathischen und unlogisch handelnden Charakteren.« (BücherTreff.de, Rez 1461075)
Nicht eindeutig/ambivalent (10.3.3)	Wird ein Aspekt (z. B. aus Ebene 1) nicht eindeutig gewertet?	Korpusschnitt: »Ich dachte, hier geht es um eine witzige Geschichte über vier frustrierte Fußballerfrauen und ihrem gemeinsamen Rachefeldzug, aber durch Annas Schicksal ist das alles so nebensächlich geworden. Deshalb kann ich mich nicht entscheiden, ob mir das Buch gefallen hat, es hat einfach einen faden Nachgeschmack hinterlassen, bzw. nicht dem entsprochen, was ich erwartet habe.« (BücherTreff.de, Rez 10714)
Neutral (10.3.4)	Wird ein Aspekt (z. B. aus Ebene 1) neutral gewertet?	Korpusschnitt: »Es lohnt sich einfach nur mal da gewesen zu sein, egal wer oder was grade [sic!] ›ausgestellt‹ wird.« (Tripadvisor, Rez 6336)

Abb. 1: Ausschnitt aus dem Kategoriensystem / Tagset zur Codierung rezensiver Texte des Hildesheimer Forschungsprojektes Rez@Kultur. [Aus: Graf et al. 2021, S. 401]

Um die Häufigkeit bestimmter Kategorien in den rezensiven Texten weiter zu untersuchen, wurde eine manuelle Codierung von mehreren Texten durchgeführt. Das händische Codieren der Daten mit den oben genannten Kategorien der Wertung und der Nichteindeutigkeit von Wertungen (sowie weiteren Kategorien) erfolgte in zwei Durchgängen in den Jahren 2018 und 2019. Im Jahr 2018 wurden 140 Buchrezensionen aus dem Amazon-Rezensionsdatensatz annotiert. Im Jahr 2019 wurden den Codierenden ca. 400 Buchrezensionen des digitalen Buchforums BücherTreff.de zur Verfügung gestellt. Die optimale Anzahl von Rezensionen wurde auf der Grundlage ihrer Länge und der durchschnittlichen Lesegeschwindigkeit eines ausgebildeten Erwachsenen gewählt, da für das Annotationsprojekt ein Zeitlimit vorgegeben war. Der Schwerpunkt der Annotation des Amazon-Datensatzes lag auf der Validierung des Kodierungssystems. Deshalb wurden die Rezensionen aus dem Amazon-Datensatz nach dem Zufallsprinzip ausgewählt, wobei das einzige Kriterium die Länge der Rezension war. Für den zweiten Datensatz wurden die gesamten Rezensionenthreads zu Büchern verschiedener Genres ausgewählt. Dabei wurden die Länge eines Threads und die Länge der ersten Rezension und der nachfolgenden Antworten darauf berücksichtigt. Jeder Beitrag innerhalb eines Threads wurde als eine Rezension betrachtet. In beiden Sets waren Rezensionen zu verschiedenen medialen Präsentationsformen von Büchern enthalten (z. B. zu gedruckten Büchern, E-Books und Hörbüchern) sowie zu unterschiedlichen literarischen Genres und Gattungen (z. B. zu Sachbüchern ebenso wie zu Fantasy-, Horror- oder Gegenwartsliteratur).²⁷

[12]

²⁷ Bei den Codierenden handelte es sich um Studierende mit grundlegender, allerdings variierender Expertise zu Rezensionstexten aus verschiedenen Disziplinen, unter anderem aus den Studiengängen Internationale Kommunikation und Übersetzen, Kulturwissenschaften und künstlerische Praxis, Literarisches Schreiben und Kulturjournalismus und Internationale Fachkommunikation an der Universität Hildesheim.

Die Aufgabe entsprechend des Codierleitfadens bestand darin, Inhalte oder Sprachhandlungen und -haltungen der Rezensionen mit Kategorien aus dem Kategoriensystem zu markieren. Das Kategoriensystem wurde multidimensional angelegt, sodass in der Regel eine Referenzkategorie der Ebene 1 (z. B. Bezug auf den Stil des Werkes) mit einer Sprechhandlung der Ebene 2 gemeinsam annotiert werden konnte (z. B. positive Wertung). Die Mehrfach-Codierung betraf als allgemeines Prinzip auch die Unterkategorien. Den Codierenden lag jeweils die gesamte Rezension vor, wobei sie selbst bestimmen konnten, welche Kategorien sie verwendeten und welche Textabschnitte sie jeweils damit markierten. Ein codierter Textabschnitt kann daher ein Wort oder auch einen ganzen Absatz bis hin zur gesamten Rezension umfassen. Umgekehrt kann eine Rezension bzw. ein Textabschnitt nur sehr selektiv oder (theoretisch) gar nicht mit Codes versehen werden.

[13]

Auch bei der Anzahl der Kategorien, die sie zur Codierung des Textes verwenden konnten, waren die Codierenden nicht eingeschränkt. Dass Mehrfachcodierungen möglich waren, ist für die Auswertung der Nichteindeutigkeit von Wertungen insofern relevant, als diese somit auf zwei Arten markiert werden können: Zum einen durch die Markierung mit der Kategorie »nicht eindeutige Wertung«, zum anderen durch die Zweifachcodierung einer Textstelle mit den beiden Kategorien »positive Wertung« und »negative Wertung«. Bei der nachfolgenden Auswertung der Codierungen hinsichtlich der Nichteindeutigkeit von Wertungen wurde zudem aus pragmatischen Gründen die Kategorie »neutrale Wertung« ebenfalls als Form der Nichteindeutigkeit angesehen und somit unter diese Ergebnisklasse subsummiert. Als »eindeutig« wurden also nur klar positive oder negative Wertungshandlung aufgefasst, während neutrale Wertungen per se als nichteindeutig verstanden wurden.

[14]

Für die Eindeutigkeit von Wertungen bedeuten die verfügbaren Codes aus dem Kategoriensystem zudem, dass nicht explizit markiert werden konnte, wenn Bewertungen eindeutig vorgenommen wurden. »Eindeutigkeit der Wertung« stellte stattdessen den Default-Fall dar. Die Fälle eindeutiger Wertung lassen sich zwar ex negativo extrahieren, allerdings kann es durch die Annahme eines »Normalfalles« der Eindeutigkeit zu einer Unterrepräsentation von nichteindeutigen Bewertungen gekommen sein, die nur leicht ambig sind. Dies kann beispielsweise vorkommen, wenn die Nichteindeutigkeit einer Wertung unterhalb der bewussten Wahrnehmungsschwelle der Codierenden liegt (man denke an relativierende Ausdrücke wie »ziemlich«) und die Signale zur Anwendung der zusätzlichen Kategorie »nicht eindeutige Wertung« nicht ausgelöst werden. Eine gewisse Korrektur erfuhren derartige Tendenzen durch die Einführung eines Goldstandards bei der Codierung: So wurde die Entscheidung, ob ein Textabschnitt zu einer der Kategorien gehörte, entsprechend der Mehrheit der Codierungen getroffen.

[15]

Das gewählte Vorgehen der induktiven Kategorienbildung, die weitestgehend alltagssprachlich-intuitive Anwendung der Kategorien durch die Codierenden und die Anwendung des Goldstandards der Mehrheitsmeinung hat viele Vorteile und entspricht dem Vorgehen der *Grounded Theory Methodology*.²⁸ Eine Konsequenz dieses Vorgehens ist allerdings auch, dass die Konzepte der Nichteindeutigkeit und der Neutralität nicht a priori theoretisch begründet werden können. Das Urteil über die Nichteindeutigkeit von rezensiven Texten hängt hier also entscheidend davon ab, was die Codierenden in der Mehrheit als nichteindeutig oder neutral empfinden. Im Anschluss an die quantitative Auswertung der Codierungen erfolgt daher weiter unten die Gegenprobe anhand von Sentimentwortlisten sowie eine qualitative Sichtung der codierten Textstellen. Dadurch wird es möglich, zu präzisieren, was von den Codierenden als »nicht-eindeutige Wertung« verstanden wurde und welche blinden Flecken im Vergleich zur oben eingeführten Arbeitsdefinition die Einschätzungen der Codierenden haben könnten.

[16]

²⁸ Vgl. Glaser / Strauss 2010 bzw. Strauss / Corbin 1996.

4. Quantitative Befunde

Die Auswertung der Codierungen zu Eindeutigkeit und Nichteindeutigkeit in Rezensionen auf Amazon und BücherTreff.de erfolgte schrittweise und hypothesengeleitet: Zuerst wurden alle Codes ausgewertet, unabhängig davon, ob sie sich auf dieselbe oder verschiedene Rezensionen bezogen (nachfolgend bezeichnet als Auswertung *auf Aussageebene*). Zweitens wurden die Codierungen auf Rezensionsebene untersucht, wobei anhand der Rezensions-ID die verschiedenen Codierungen jeweils den Rezensionen zugeordnet wurden, auf die sie sich bezogen (nachfolgend Auswertung *auf Rezensionsebene* genannt). Die Ergebnisse der Auswertung auf Rezensionsebene wurden zudem mit einem zweiten Messinstrument geprüft, indem die annotierten Rezensionen mit Sentimentwortlisten danach durchsucht wurden, ob nur positive oder nur negative Begriffe in Rezensionen (etwa nur Wörter aus den positiven Wortlisten oder nur aus den negativen Wortlisten) ähnlich selten vorkommen wie nichteindeutige Wertungen (Rezensionen, die Wörter aus beiden Listen enthielten). In einem dritten Schritt wurde danach gefragt, welcher Zusammenhang zwischen der Eindeutigkeit von Rezensionen und der pro Rezension vergebenen Anzahl von »Sternen« als einer typischen Form der numerischen Wertung im digitalen Raum besteht. Hierzu fand also eine Auswertung der Codierungen *auf Ebene der Sternbewertungen* statt.

[17]

4.1 Eindeutigkeit auf Aussageebene

Die deutliche Mehrzahl der codierten Bewertungen sowohl im Amazon- als auch im BücherTreff.de-Korpus fällt eindeutig positiv aus. Insofern bestätigen sich hier also Annahmen, die davon ausgehen, dass Online-Rezensionen typischerweise positiv und in diesem positiven Urteil auch sehr eindeutig ausfallen. Die Ergebnisse der beiden Plattformen unterscheiden sich leicht (statistisch allerdings nicht signifikant) voneinander. Auf Amazon waren insgesamt 170 Codierungen eindeutig positiv (70,5 % der insgesamt annotierten Codierungen), 39 eindeutig negativ (16,2 %) und 32 Codierungen nichteindeutig oder neutral (13,3 %). Auf BücherTreff.de fällt das Verhältnis noch deutlicher zugunsten eindeutig positiver Bewertungen aus: Mit 279 Codierungen waren über 89,4 % eindeutig positiv, während lediglich 11 (3,5 %) eindeutig negativ ausfielen. Der prozentuale Anteil an nichteindeutigen / neutralen Bewertungen ist im Vergleich zu den Codierungen des Amazon-Korpus allerdings mit 22 Codierungen und somit 7 % ebenfalls geringer.

[18]

Bei der Interpretation der Zahlen muss allerdings berücksichtigt werden, dass in den beiden Korpora jeweils unterschiedlich hohe Anteile an Rezensionen überhaupt mit Codes aus dem relevanten Bereich der eindeutigen / nichteindeutigen Bewertungen codiert wurden. Dadurch relativiert sich die hohe Prozentzahl mit Blick auf die Gesamtheit der rezensiven Texte – und zwar bei BücherTreff.de stärker als bei Amazon. Auf Amazon wurden 104 von 140 Rezensionen mit einem der relevanten Codes annotiert (d. h. 74 %). Im Korpus von BücherTreff.de wurden von insgesamt 378 Rezensionen nur 142 mit Codierungen der Wertung versehen (d. h. 37,5 %). Vor diesem Hintergrund lässt sich das Ergebnis auf Aussageebene wie folgt konkretisieren: Circa drei Viertel der Online-Rezensionen auf Amazon enthalten laut der Einschätzung der Codierenden Bewertungen. Diese 104 Amazon-Rezensionen wurden mit 241 Codierungen aus dem Bereich der eindeutigen / nichteindeutigen Bewertung versehen (also etwas mehr als zwei Codierungen pro Rezension), wobei der überwiegende Anteil auf die eindeutig positiven Wertungen entfiel. Auf der Leseplattform BücherTreff.de wurden lediglich ca. 40 Prozent der rezensiven Texte überhaupt mit einer Wertungscodierung versehen, was darauf schließen lässt, dass ein Großteil der Texte nicht wertend vorgeht. Das Kommunikationsziel auf BücherTreff.de als einem »Bücherforum« ist also möglicherweise ein anderes als die wertende und auf eine Kaufentscheidung ausgerichtete Empfehlung auf Amazon. In den 142 annotierten Rezensionen aus dem Korpus von BücherTreff.de wurden 312 Aussagen (also im Durchschnitt ca. zwei Codierungen pro Rezension) mit Codierungen aus dem Spektrum der eindeutigen / nichteindeutigen Wertung versehen. Der Anteil der eindeutig positiven Wertungen war hier ebenfalls dominant und im Vergleich mit der Amazon-Stichprobe höher.

[19]

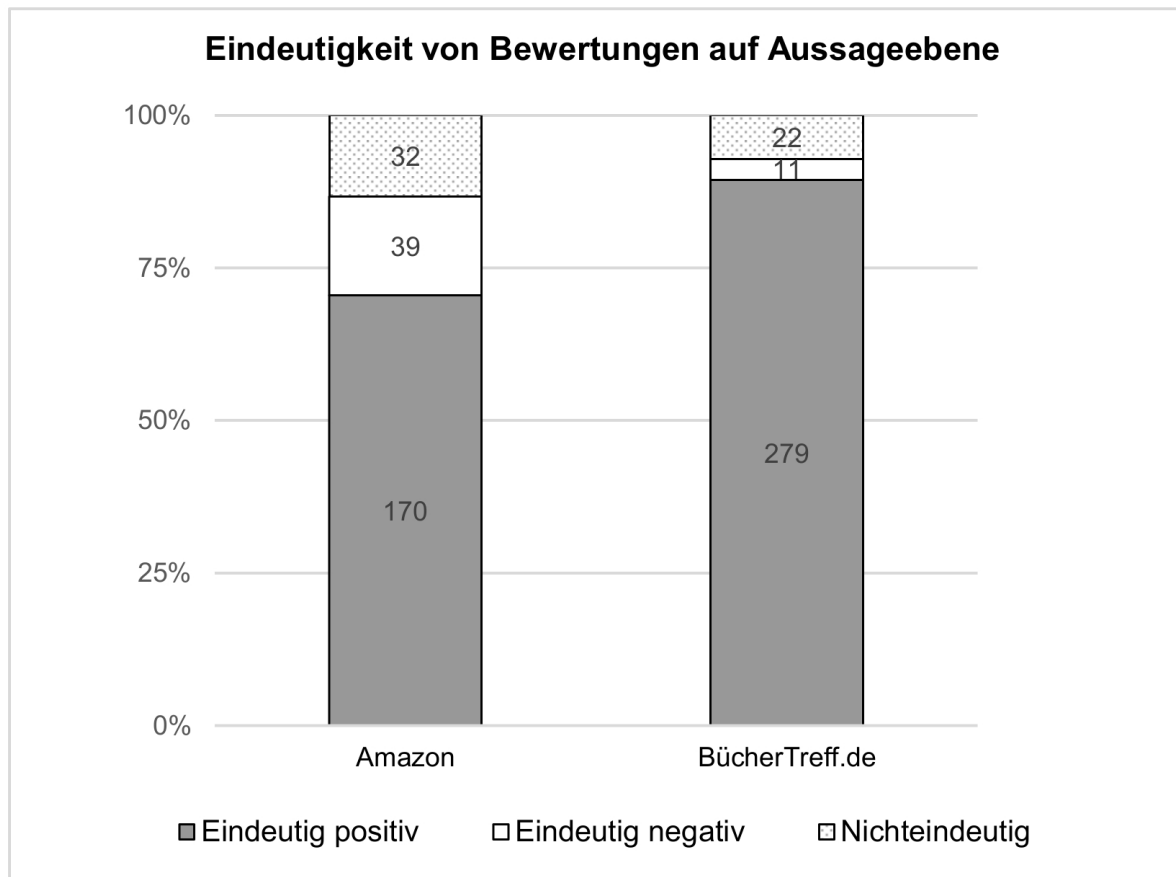


Abb. 2: Anteile als nichteindeutig und eindeutig codierter Textabschnitte an den mit diesen Codes versehenen Rezensionen auf Amazon und BücherTreff.de in absoluten und relativen Zahlen.

Auf der Basis dieser Erkenntnisse lässt sich nun weiter fragen, wie sich derartige Wertungen auf Aussageebene inhaltlich auf die codierten Rezensionen verteilen und ob durch das gemeinsame Auftreten von eindeutig positiven und eindeutig negativen Wertungen innerhalb *einer* Rezension (und damit mutmaßlich in Bezug auf denselben Gegenstand) in der Summe mehr nichteindeutige Rezensionen vorliegen, als es die Ergebnisse auf Aussageebene nahelegen. Dieser Fragestellung widmet sich daher der nächste Analyseschritt.

[20]

4.2 Eindeutigkeit auf Rezensionsebene

Durch die Auswertung der Codierungen je Rezension ergibt sich ein etwas anderes Bild als auf der Aussageebene. Amazon und BücherTreff.de enthalten demzufolge prozentual höhere Anteile nichteindeutiger Rezensionen als bei der Auswertung auf Aussageebene, während der Anteil der eindeutig positiven Rezensionen leicht sinkt. Auf Rezensionsebene ist demnach durch das gleichzeitige Auftreten von negativen und positiven Aussagen der Grad von Nicht-Eindeutigkeit der Bewertungen erhöht. Oder anders ausgedrückt: Einige eindeutig positive Wertungen werden noch innerhalb derselben Rezensionen mit eindeutig negativen Wertungsaussagen konterkariert. In diesen Fällen wird ein differenziertes Urteil dadurch herbeigeführt, dass sowohl positive als auch negative Aspekte eines Werkes thematisiert werden (vgl. hierzu die Analyse einer qualitativen Stichprobe dieser als nichteindeutig klassifizierten Rezensionen in Kapitel 5). Allerdings bleiben diese differenzierenden Rezensionen mit 26,9 % auf Amazon und 15,5 % auf

[21]

BücherTreff.de weiterhin zahlenmäßig deutlich hinter den eindeutig positiven Rezensionen mit 69 % (72 Rezensionen) auf Amazon und 83 % (118 Rezensionen) auf BücherTreff.de zurück. Auch hier erwiesen sich die Unterschiede zwischen den beiden Plattformen als statistisch nicht signifikant.

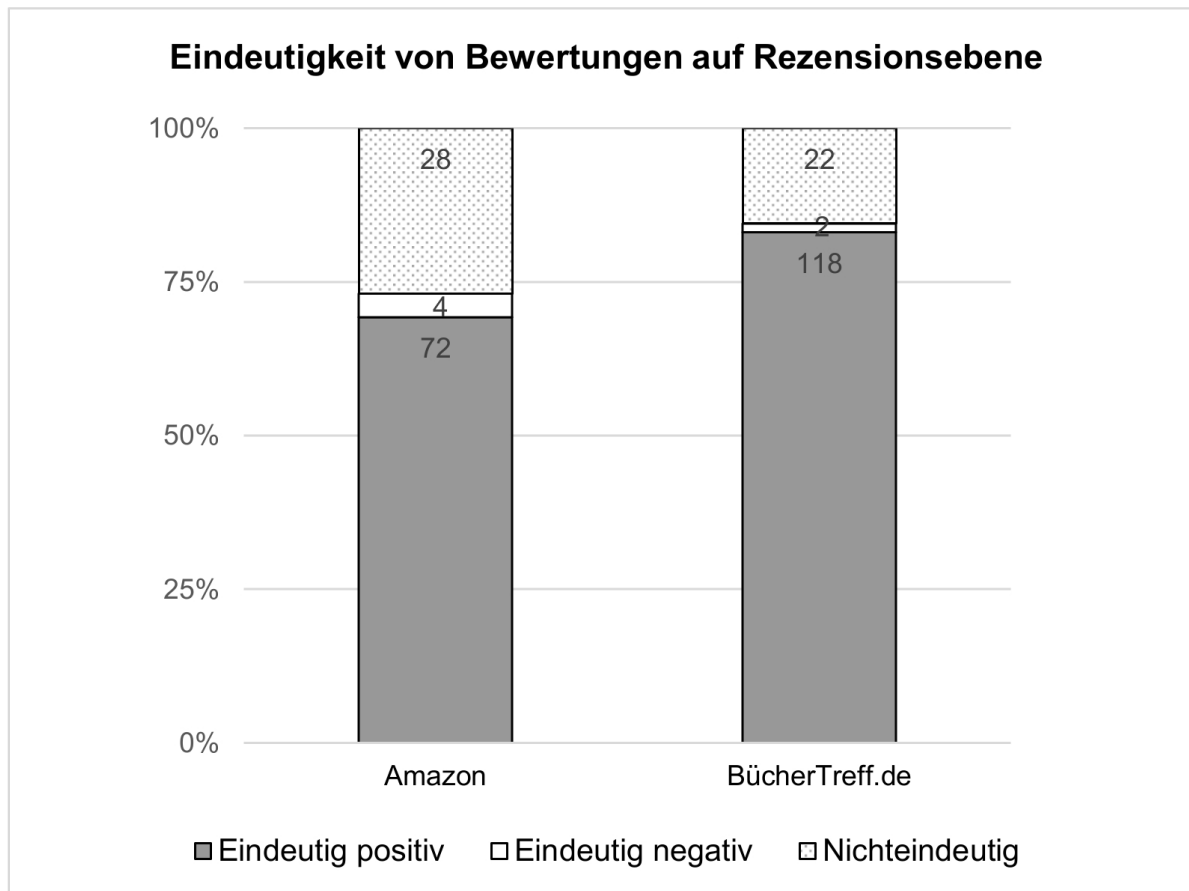


Abb. 3: Anteile nichteindeutig und eindeutig wertender Rezensionen innerhalb der Gesamtmenge an mit diesen Codes versehenen Rezensionen auf Amazon und BücherTreff.de in absoluten und relativen Zahlen.

Insgesamt muss auch hier wieder berücksichtigt werden, dass der Gesamtanteil der mit Codes der Wertung versehenen Rezensionen auf Amazon deutlich höher ist als jener auf BücherTreff.de. Es relativiert sich also erneut die deutliche Dominanz eindeutig positiver Rezensionen besonders auf BücherTreff.de, wo immerhin 62,5 % der Rezensionen nicht mit Codes der Wertung annotiert wurden. Die Mehrheit der Rezensionen, die keine Wertungs-Codes hatten, bezogen sich auf den Inhalt der Geschichte oder die Metadaten zum Buch: Verlag, Seitenzahl, Namen der Übersetzer.

[22]

Im Falle des Amazon-Datensatzes entfallen gegenüber der nur etwas geringeren Anzahl der eindeutig positiven Rezensionen (im Vergleich zu den eindeutig positiven Aussagen desselben Korpus) die eindeutig negativen Aussagen fast vollständig auf die nichteindeutigen Rezensionen. Der Anteil eindeutig negativer Rezensionen beträgt auf Amazon lediglich 3,8 % (4 Rezensionen), während immerhin über 16 Prozent der wertenden Aussagen eindeutig negativ sind. Dieser Effekt ist auch im Datensatz von BücherTreff.de zu beobachten, allerdings etwas schwächer ausgeprägt. Die ungleiche Verteilung von eindeutig negativen Aussagen auf die Rezensionen legt nahe, dass nichteindeutige Rezensionen der bevorzugte Ort kritischer Äußerungen auf Amazon und BücherTreff.de sind, während vollständig negative Rezensionen sehr selten vorkommen. Entweder finden Rezensent*innen die Bücher tatsächlich selten ausschließlich schlecht oder sie beschreiben »schlechte« Bücher dennoch nicht eindeutig negativ. Für beide Erklärungen gibt es gute Argumente: So liegt bereits der Kaufentscheidung eine Wertungshandlung zugrunde, die in der Regel darauf

[23]

zielt, die nach den eigenen Kriterien »lohnenswerten« Bücher zu lesen und zu kaufen und nicht solche, die den eigenen Präferenzen gar nicht entsprechen (wobei es solche Lektüreentscheidungen natürlich auch gibt). Insofern würden nur echte »Fehlkäufe« eindeutig negative Rezensionen erzeugen, wobei selbst in solchen Fällen möglicherweise ein Mechanismus der Selbstbestätigung und des Identitätsmanagements dazu führen könnte, dass auch eindeutig enttäuschende Kaufentscheidungen retrospektiv durch eine nichteindeutige Rezension aufgewertet werden. Denkbar ist auch, dass formatspezifische Konventionen (wie das Ratingsystem für besonders »nützliche« Rezensionen auf Amazon) dazu führen, dass schlechte Bücher nicht eindeutig negativ besprochen werden, sondern dass eine Ausgewogenheit bzw. Differenziertheit im Urteil angestrebt wird. Belegt ist zudem, dass insbesondere Rezensionen auf Kaufplattformen primär als Praktiken der Empfehlung aufgefasst werden, in denen weniger der literarisch-kritische Diskurs als Vorbild fungiert – indem Werke beispielsweise mit einem pädagogischen Impetus auch gänzlich »verrissen« werden dürfen –, sondern vielmehr eine Weitergabe eigener – bestenfalls geglückter – Kauf- und Leseentscheidungen angestrebt wird.²⁹ Auch diese Zielsetzung könnte eine zu positiven und ausgewogenen Rezensionen hin tendierende Wertungspraxis begründen.

4.3 Eindeutigkeit numerischer Wertungen

Auf vielen Online-Handels- und Social-Reading-Plattformen ist es verbreitete Praxis, dem rezensiven Text ein numerisches Urteilsäquivalent in Form einer Anzahl von zumeist ein bis fünf möglichen Sternen beizufügen.³⁰ [24] Holger Kellermann und Gabriele Mehling konnten belegen, dass in Buchrezensionen auf Amazon eine hohe Anzahl von Sternen – im Durchschnitt 4,1 Sterne – vergeben werden.³¹ Mit 60 % Anteil an der Gesamtheit der von Kellermann / Mehling untersuchten Rezensionen überwogen dabei die 5-Sterne-Rezensionen, wie auch in der Stichprobe von Andrea Bachmann-Stein aus dem Jahr 2015.³²

Sternebewertungen stellen dabei eine numerische »Quintessenz der Rezension«³³ dar und sind in [25] die algorithmischen Aufmerksamkeitslogiken vieler Plattformen eingebunden. Numerisch-grafische Bewertungsindizes sind aber auch ein Mittel der Komplexitätsreduktion und der »Vereindeutigung«. So beschreibt Steffen Mau in seinen Überlegungen *Das metrische Wir*: »Zahlen vermitteln also Präzision, Eineindeutigkeit, Vereinfachung, Nachprüfbarkeit und Neutralität.«³⁴ Und Johannes Spengler konkretisiert diesen Zusammenhang in Bezug auf digitale Rezensionspraktiken und deren Funktion als Selbstvermessung wie folgt: »Unübersichtliche Gegenstände und Sachverhalte wie etwa die Bewertungen komplexer Kulturprodukte werden durch Zahlen in übersichtliche Ordnungen und Hierarchien übersetzt: [...] Quantifizierung ist Reduktion mit dem Ziel, Ordnung herzustellen.«³⁵

Dass Quantifizierungen zur »Vereindeutigung« von Urteilen beitragen, muss allerdings in Bezug auf [26] Sternebewertungen auf Amazon differenziert werden: Erstens ist zwischen Gemeintem und Gezeigtem zu unterscheiden. Eindeutigkeit und empirische Faktizität werden durch die Sterne zwar suggeriert, allerdings gibt es weder eine einheitliche Skala noch einheitliche Bewertungskriterien. Die universelle Vergleichbarkeit der Produkte und Bücher wird also lediglich suggeriert, aber das *tertium comparationis* (im Hinblick *worauf* die Bücher also vergleichbar sein sollen) ist nicht klar und unterscheidet sich zudem von Produkt(art) zu Produkt(art): Ein 5-Sterne-Buch kann von Nutzer*innen positiv bewertet werden, weil es ihre Erwartungen an einen Genre-Roman erfüllt, während ebendieses Buch nach Kriterien der literarischen Qualität von Literaturkritiker*innen nicht mehr als einen Stern verdient hätte. In den Rezensionen wird dabei oft deutlich,

²⁹ Vgl. Rehfeldt 2017, S. 248.

³⁰ Vgl. hierzu die jüngste Forschung zur Quantifizierung des Selbst und des Sozialen, u. a. Mämecke 2021; Mau 2017.

³¹ Vgl. Kellermann / Mehling 2017, S. 188.

³² Vgl. Bachmann-Stein 2015, S. 81.

³³ Vgl. Bachmann-Stein 2015, S. 80.

³⁴ Mau 2017, S. 12.

³⁵ Spengler [im Erscheinen], S. 6.

dass Leser*innen in geschmacklichen Stilgemeinschaften³⁶ Bücher rezipieren und sie daher Werke anhand von gruppenspezifischen Wertungskriterien beurteilen, die aber eben nicht von allen Leser*innen oder anderen Amazon-Nutzer*innen geteilt werden müssen. Zweitens kann Eindeutigkeit auf der Ebene der numerischen Skalen zwar hergestellt werden, allerdings besteht diese Eindeutigkeit dann ausschließlich system- bzw. skalenimmanent und nicht hinsichtlich des sprachlich-inhaltlichen Urteils über das Buch: Eine 5-Sterne-Bewertung ist eindeutig eine maximal positive Bewertung und eine 1-Sterne-Bewertung ist eindeutig eine maximal negative Bewertung. Eine 3-Sterne-Bewertung ist eine eindeutig mittlere Bewertung innerhalb der Skala bis 5. Allerdings ist gerade bei den Rezensionen mit drei Sternen zu vermuten, dass hier auf der sprachlich-inhaltlichen Ebene der Rezensionen eben gerade keine Eindeutigkeit besteht. Die durch 3-Sterne suggerierte numerisch eindeutige Mittelmäßigkeit deutet also – so die Hypothese – auf inhaltliche Nichteindeutigkeit.

Die Auswertung der händischen Codierungen von Amazon- und BücherTreff.de-Rezensionen bestätigt diese Vermutung. So liegt der Anteil eindeutiger Rezensionen bei den 5-Sterne-Bewertungen auf Amazon mit 60 (85,7 %) am höchsten und bei den 3-Sterne-Bewertungen am niedrigsten mit nur einer von sieben Rezensionen (also 14,3 %). Dies korrespondiert mit der Hypothese, dass 3-Sterne-Bewertungen als mittlerer numerischer Bereich zwischen besonders positiven und negativen Wertungen eine mittlere Position einnehmen, die auf der Aussageebene der Bewertungen durch Ambivalenz und das Nebeneinander sowohl positiver als auch negativer Aspekte gekennzeichnet ist. Für Rezensionen mit einem oder zwei Sternen lassen sich aufgrund der geringen Datenmenge ($n_{\text{Amazon}} = 5$, $n_{\text{BT}} = 2$) allerdings keine verallgemeinerbaren Aussagen ableiten.

[27]

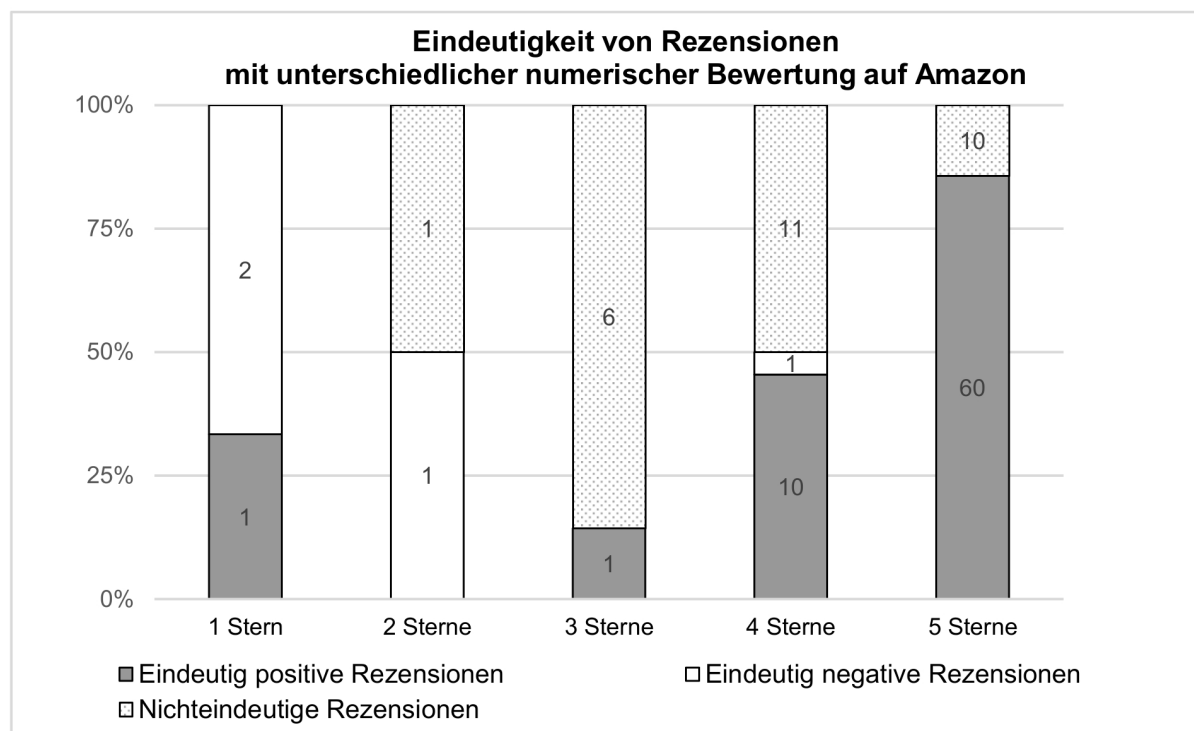


Abb. 4: Anteile nichteindeutig und eindeutig wertender Rezensionen verteilt nach der Anzahl der Sterne auf Amazon in absoluten und relativen Zahlen.

³⁶ Vgl. Baßler 2021. Baßlers Bezug auf das Konzept der Stilgemeinschaft nach Jochen Venus erweist sich in diesem Zusammenhang als vielversprechend. Baßlers recht pauschaler Kritik an der ideologischen Selbstvergewisserung der Leser*innen in solchen Gemeinschaften und an deren anti-akademischer Ablehnung literaturwissenschaftlicher Kompetenzen teilen die Autorinnen allerdings nur bedingt (o. S.). Alternativ (und neutraler) ließe sich auch von unterschiedlichen communities of practice im Anschluss an Wenger 1998 sprechen, allerdings fehlt es für eine derartige Klassifikation noch an einschlägigen Forschungsarbeiten.

Auf BücherTreff.de korrelieren Eindeutigkeit und numerische Wertung allerdings nicht so deutlich. Mögliche Erklärungen sind die grundsätzlich fakultative Sternevergabe auf der Plattform. Während auf Amazon jeder verfassten Rezension ein Sternewert zugeordnet werden *muss*, erhält diese Quantifizierungsform auf BücherTreff.de durch die Freiwilligkeit einen sekundären Status im Sinne einer Zusatzinformation. In der Mehrzahl der codierten Rezensionen auf BücherTreff.de (75 von 142, d. h. 53 %) wurden keine Sterne vergeben. Zudem ist auch hier die Anzahl von Rezensionen mit weniger als 4 Sternen recht gering, sodass eine statistische Auswertung und besonders die Verallgemeinerung nicht sinnvoll sind. Drittens ist es möglich, dass die Ausrichtung der Plattform auf den Austausch über Bücher (und nicht auf die Unterstützung einer Kaufentscheidung) dazu führt, dass das System der numerischen Markierungen etwas weniger kohärent bzw. stringent etabliert ist. So ist es auf BücherTreff.de möglicherweise legitim, dass auch die 3-Sterne-Bewertungen hauptsächlich positiv wertende Aussagen über die rezensierten Werke enthalten. [28]

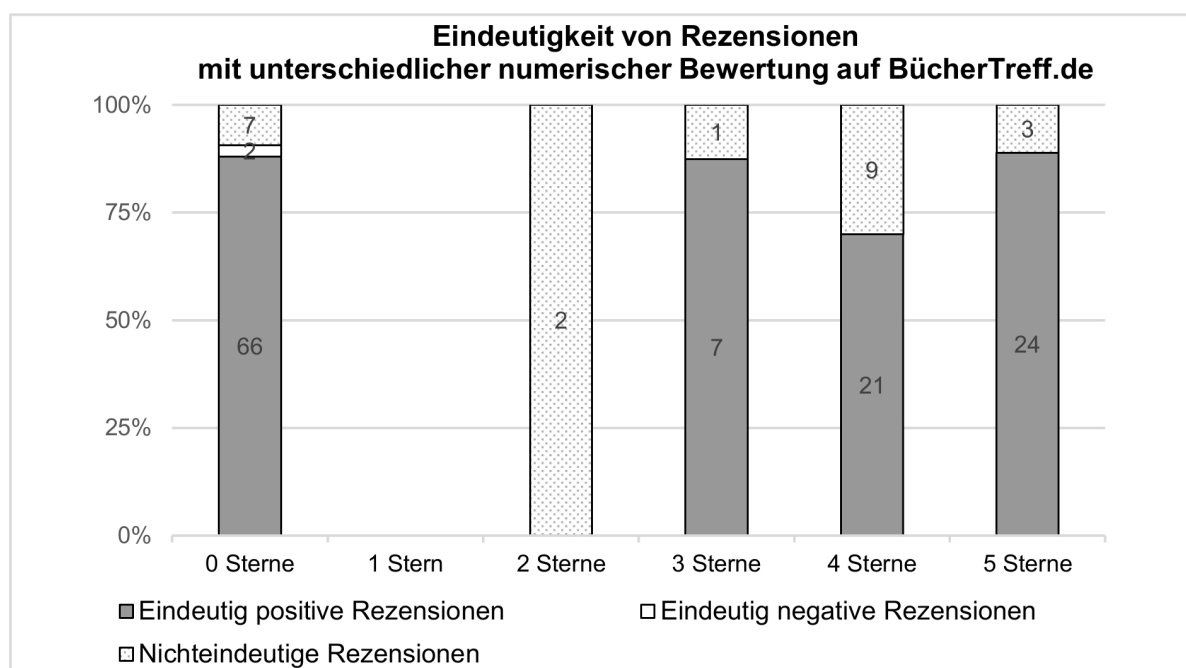


Abb. 5: Anteile nichteindeutig und eindeutig wertender Rezensionen verteilt nach der Anzahl der Sterne auf BücherTreff.de in absoluten und relativen Zahlen.

4.4 Vergleich mit Sentiment-Wortlisten

Um die Hypothese zur Verteilung von Nichteindeutigkeit in Buchrezensionen weiter zu überprüfen, wurden die Annotationen der Codierenden mit den Ergebnissen einer lexikonbasierten *Sentiment-Analyse* verglichen. Für die Sentiment-Analyse wurde das *SentiWS-Lexikon*³⁷ als Grundlage gewählt. Vorangegangene Forschungen³⁸ haben gezeigt, dass diese spezielle Ressource gut für die Analyse literarischer Texte geeignet ist. Das *SentiWS-Lexikon* zeichnet sich durch seine umfassende Erfassung von Wörtern aus, die eine positive (1.818) oder negative (1.650) Wertung ausdrücken. [29]

Bei der Sentiment-Analyse der Online-Rezensionen von Amazon und BücherTreff.de wurde zunächst die lemmatisierte Version einer Rezension mit der Liste der positiven und negativen Wörter verglichen. Anschließend wurde der Prozentsatz der positiven und negativen Wörter in der Rezension berechnet. Dabei wurden folgende Klassifizierungen vorgenommen: [30]

³⁷ Remus et al. 2010.

³⁸ Vgl. Fehle et al. 2021.

- Wenn die Rezension ausschließlich positive und nicht sentimentmarkierte, d. h. nicht auf der Sentimentwortliste verzeichnete, Wörter enthielt, wurde sie als positiv eingestuft. In diesem Fall wurde die Gesamtstimmung der Rezension als »eindeutig positiv« betrachtet.
- Wenn die Rezension ausschließlich negative und nicht sentimentmarkierte Wörter enthielt, wurde sie als »eindeutig negativ« klassifiziert.
- Wenn eine Rezension (neben nicht sentimentmarkierten Ausdrücken) sowohl positive als auch negative Wörter enthielt, wurde sie als »nichteindeutig« betrachtet.

Diese Methode basiert auf einer einfachen Wort-zu-Wort-Vergleichsbasis und berücksichtigt keine semantischen oder kontextuellen Zusammenhänge. Daher können bestimmte inhaltliche Feinheiten oder sprachliche Nuancen in der Sentiment-Analyse möglicherweise nicht erfasst werden. Dennoch bietet sie eine Möglichkeit, die grundlegende Tendenz einer Rezension zu bestimmen und erste Erkenntnisse über das affektive Feedback der Rezensent*innen zu gewinnen. Das Ergebnis dieser Analyse zeigte, welche Rezensionen »eindeutig positiv« (enthielt Wörter, die nur im positiven Lexikon gefunden wurden), »eindeutig negativ« (enthielt Wörter, die nur im negativen Lexikon gefunden wurden) oder »nichteindeutig« (enthielt Wörter aus beiden Lexika) waren.

[31]

Abbildung 6 zeigt die Verteilung der positiven, negativen und nichteindeutigen Rezensionen auf Grundlage der Ergebnisse der Sentiment-Analyse für beide Plattformen. Basierend auf dem oben beschriebenen Verfahren der Sentiment-Analyse wurden fast gleich viele Rezensionen als »nichteindeutig« wie als »eindeutig positiv« gekennzeichnet (51 Rezensionen mit ausschließlich positiven Wörtern und 49 Rezensionen mit sowohl positiven als auch negativen Wörtern für Amazon). Die Verteilung für Rezensionen von Büchertreff.de zeigt das gleiche Muster: 74 Rezensionen mit ausschließlich positiven Wörtern und 67 mit sowohl positiven als auch negativen Wörtern.

[32]

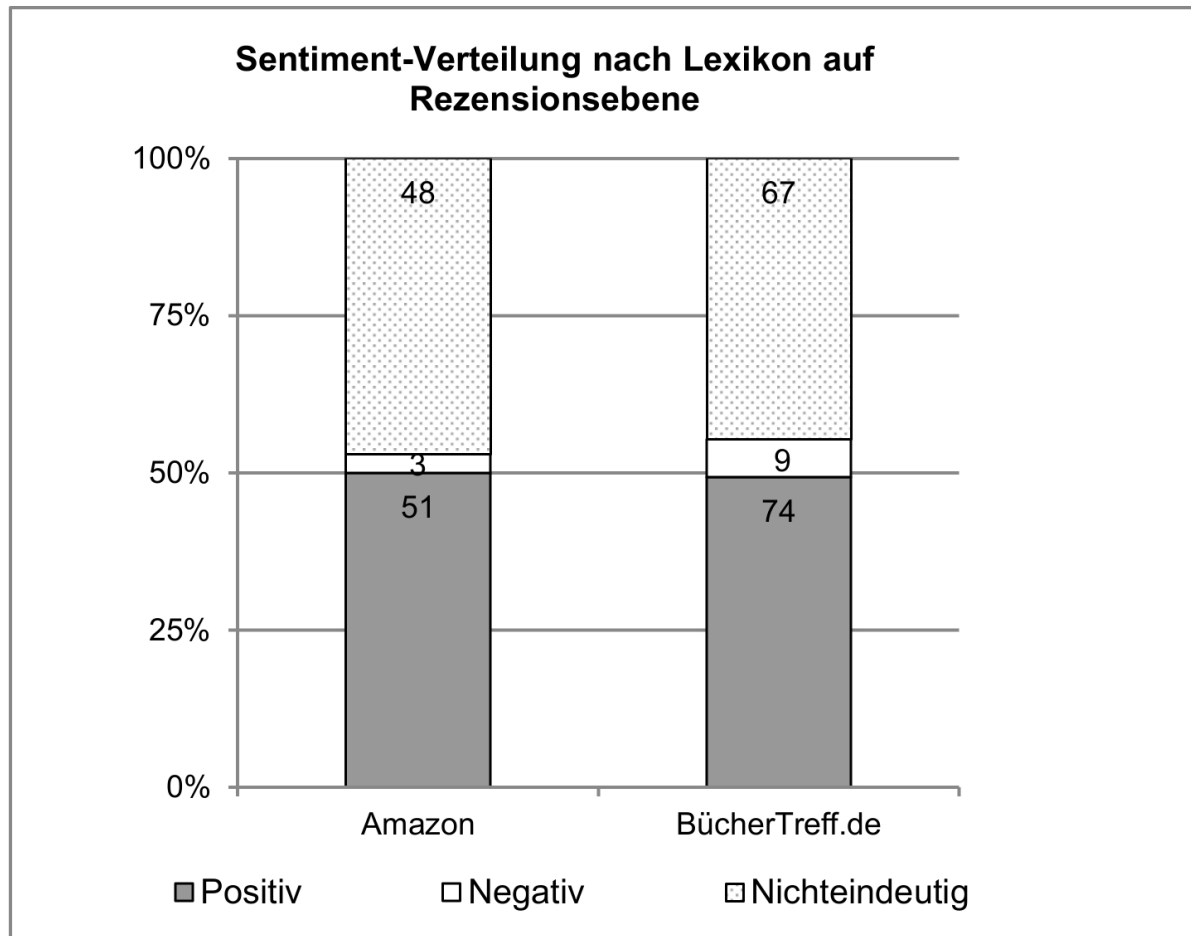


Abb. 6: Anteile von positiven, negativen und nichteindeutigen Rezensionen auf Basis der Sentiment-Analyse auf Amazon und BücherTreff.de.

Eine mögliche Erklärung für die Unterschiede zwischen manuellen Codierungen und automatischer Annotation durch die Sentiment-Analyse könnte in den Lexika selbst liegen. Oft enthalten diese Listen Wörter, die im Kontext einer Rezension eine andere Bedeutung haben können. Ein Beispiel hierfür wäre das Wort »Hammer«, das in der Sentimentwortliste negativ markiert ist. Es kann in einer Rezension allerdings positiv verwendet werden, um etwas als außergewöhnlich oder hervorragend zu beschreiben. Diese Unterschiede in der Bedeutung können zu Ungenauigkeiten oder Fehlinterpretationen in der Sentiment-Analyse führen. Aus der Differenz zwischen den händisch codierten Rezensionen (die häufiger eindeutig positiv waren) und dem höheren Anteil nichteindeutiger Rezensionen bei der Sentiment-Analyse lässt sich also schließen, dass Online-Rezensionen nicht unbedingt in einem einfachen, wortwörtlichen Sprachstil verfasst sind, sondern dass Formen der uneigentlichen Sprache (wie das »Hammer«-Beispiel) oder auch Verneinungen verbreitet sind (denn zwei negative Sentimentwörter können in Kombination eine positive Bedeutung aufweisen), welche das Ergebnis der automatisierten Annotation hin zur Nichteindeutigkeit beeinflussen. Umgekehrt könnte man auch davon ausgehen, dass menschliche Codierende die vermeintlichen Widersprüche, die sich für den Computer aus dem gleichzeitigen Auftreten von Sentimentwörtern konträrer Register ergeben, nicht als nichteindeutig wahrnehmen, weil sie diese durch Inferenzen, Kontext- und Weltwissen »vereindeutigen« können.

5. Qualitative Befunde

Die statistische Annäherung an die Daten legt nahe, dass sowohl in Rezensionen auf Amazon als auch auf BücherTreff.de eindeutig positive Wertungen überwiegen – sofern Wertungen vorliegen. Zweitens treten eindeutig negative Aussagen selten in ausschließlich negativ wertenden Rezensionen auf, sondern eher eingebettet in differenzierende und damit nichteindeutige Rezensionen, in denen sowohl positive als auch negative Aspekte eines Werkes thematisiert werden. Zuletzt scheint die Eindeutigkeit dort, wo die Größe des Datensatzes eine solche Auswertung zulässt, mit der Anzahl der vergebenen Sterne zu korrelieren: Mittlere numerische Werte (3 Sterne) gehen häufiger mit nichteindeutigen Rezensionen einher, während die Randbereiche der Skala (also minimale und maximale Sternestahlen) häufiger auf eindeutige Rezensionen (positive oder negative) entfallen. Allen statistischen Ergebnissen liegt dabei eine von den menschlichen Codierenden intuitiv und alltagssprachlich angewendete Definition von Eindeutigkeit bzw. Nichteindeutigkeit zugrunde, die sich im Abgleich mit den Sentimentwortlisten bereits als eher restriktiv bzw. konservativ erwiesen hat: Mithilfe der Sentimentwortlisten konnten höhere Werte an nichteindeutigen Rezensionen erfasst werden³⁹, was darauf hindeutet, dass die menschlichen Codierenden in der Tendenz nichteindeutige Wertungen nicht immer annotiert haben und / oder dass sie ko- und kontextuell basierte Eindeutigkeit besser erkennen als es die automatische Auswertung der Sentimentwörter vermag. Welche Vorstellung von Nichteindeutigkeit genau dem Codierungsverfahren zugrunde lag, ist bisher noch unklar und soll durch eine qualitative Stichprobe der codierten Rezensionen beleuchtet werden.

[34]

Die vergebenen Sterne bildeten dabei in Kombination mit der quantitativen Auswertung der händischen Codierungen die Grundlage für die Auswahl der Texte, die im *Close-Reading*-Verfahren⁴⁰ analysiert wurden. In beiden Teilkorpora wurden jeweils Rezensionen qualitativ hinsichtlich ihrer Eindeutigkeit betrachtet, die numerische Extremwerte erzielt hatten (d. h. die Rezensent*innen hatten jeweils fünf oder einen Stern/e an das Buch vergeben) und die gleichzeitig von den Codierenden entweder als eindeutig positiv oder als nichteindeutig codiert wurden. Derselbe Suchvorgang wurde auf die 1- und die 3-Sterne-Rezensionen aus dem Amazon- und dem BücherTreff.de-Korpus angewendet. Qualitativ ausgewertet wurden somit:

[35]

- eindeutig positive 5-Sterne-Wertungen;
- nichteindeutige 5-Sterne-Wertungen;
- eindeutig negative 1- und 2-Sterne-Wertungen;
- nichteindeutige 1- und 2-Sterne-Wertungen;
- 3-Sterne-Bewertungen.

Im Gegensatz zu der intuitiven Codierung von Eindeutigkeit im Fall der quantitativen Auswertung lag der qualitativen Analyse die weiter oben eingeführte Arbeitsdefinition von Eindeutigkeit zugrunde. Als »eindeutige Wertung« wurde demnach betrachtet, wenn in dem jeweils vorliegenden sprachlich-pragmatischen Kontext einem Referenzobjekt ein attributiver Wert eindeutig zugeordnet wird, keine anderen Lesarten existieren und auch keine Interpretationsspielräume eröffnet werden, die das gefällte Werturteil in Frage stellen.

[36]

5.1 Nichteindeutigkeit in 5-Sterne-Rezensionen

Ein erster qualitativer Befund besteht in der Erkenntnis, dass der Spielraum für Nichteindeutigkeiten selbst dann noch groß ist, wenn Rezensionen numerische Maximalwerte aufweisen und von Codierenden als eindeutig positiv (mit mindestens je vier eindeutig positiven Codes, $n = 7$) wahrgenommen wurden.

[37]

³⁹ Es ist sehr wichtig hier zu erwähnen, dass die Lexika ursprünglich nicht zusätzlich auf die Aufgabe der Annotation von Buchrezensionen zugeschnitten waren. Ihnen fehlte z. B. die umgangssprachliche Nutzung von Wörtern wie »Hammer«. Außerdem wurde die Rezension von uns automatisch als nichteindeutig eingestuft, sobald zwei Wörter aus verschiedenen polaren Lexika in der Rezension gefunden wurden.

⁴⁰ Vgl. Smith 2016.

Umgekehrt können nichteindeutig codierte Rezensionen durchaus eindeutig in ihrer numerischen Wertung sein ($n = 4$). In Beispiel 1 wird die Neuerzählung der Grimm'schen Märchen von Philip Pullman gelobt: Die beigefügten Kommentare machten »dieses Märchenbuch definitiv zu etwas Besonderem«. Es sei »sehr empfehlenswert« und »für Märchen [sei] man einfach nie zu alt.« Die Rezension wurde von den Codierenden ausschließlich und sogar mehrfach mit dem Code »eindeutig positive Wertung« versehen. Dennoch handelt es sich insgesamt nicht um einen reinen »Lobgesang«. Denn die Rezensent*in erzeugt bewusst einen Spannungsbogen und damit eine kurzzeitige Nichteindeutigkeit, indem sie in einem ersten Schritt die Vorhersehbarkeit und den Schematismus der Gattung der Märchen kritisiert:

»Prinzessinnen sind schön, Wälder dunkel, Stiefmütter böse. Damit hat sich das.«

Allerdings fungiert diese pejorative Charakterisierung hier nur als rhetorische Form der Nichteindeutigkeit, welche die positive Bewertung des Werkes nur umso deutlicher hervorheben soll. Ein Vorurteil wird nur deshalb angeführt, um es im nächsten Schritt zu widerlegen.

(1)

[38]

»[...] Prinzessinnen sind schön, Wälder dunkel, Stiefmütter böse. Damit hat sich das. Und trotzdem lieben wir Märchen. [...] Gerade diese Kommentare machen dieses Märchenbuch definitiv zu etwas Besonderem ein Buch zum Lesen, Vorlesen und Wiederlesen. Sehr empfehlenswert. Für Märchen ist man einfach nie zu alt.«

(Strandläuferin / BücherTreff.de: Rezension zu *Grimms Märchen* von Philip Pullman, 5 Sterne. 21.09.2013. HTML. [\[online\]](#))

Die nachfolgende Beispielrezension wurde von den Codierenden in der Mehrzahl mit der Kategorie »eindeutig positive Wertung« kodiert. Allerdings erhielt die Rezension auch zwei Codierungen für eindeutig negative Aussagen. Dass mit der Rezension insgesamt fünf Sterne für das rezensierte Werk *Ich. Darf. Nicht. Schlafen.* von S. J. Watson vergeben wurden, lässt sich damit erklären, dass sich die negativen Codierungen auf eine Aussage zum englischsprachigen Titel (*Before I Go To Sleep*) beziehen. In diesem Fall liegt also eine Gewichtung der konfligierenden Wertungshandlungen vor, welche das Gesamtergebnis der Wertung zu einer Seite, der eindeutig positiven, kippen lässt:

[39]

(2)

[40]

»[...] Der Titel der Originalausgabe gefällt mir leider nicht so gut wie der deutsche Titel »Ich. Darf. Nicht. Schlafen.« [...] Insgesamt gibt es von mir fünf von fünf Sternen: Gute Idee, guter Aufbau, guter Schreibstil, gute Charaktere und am Ende ein zufriedener Leser.«

(Malin Jo / Amazon: Rezension zu *Ich. Darf. Nicht. Schlafen.* von S. J. Watson, 5 Sterne. 10.11.2014. HTML. [\[online\]](#))

Die Auswertung der numerisch maximal positiv bewerteten Rezensionen verdeutlicht zwei wichtige Aspekte der Nichteindeutigkeit im Kontext literarischer Wertung: Erstens belegen die Beispiele das Auftreten dynamischer und rhetorischer Nichteindeutigkeit, die als *Prokatalepsis* im Rezensionsverlauf wieder aufgelöst werden kann: Im Märchen-Beispiel (1) ist nicht klar, ob die Rezensent*in der Aussage, Märchen seien langweilig, wirklich zustimmt und dies nur im Beispielfall nicht gegeben sieht, oder ob sie grundsätzlich anderer Meinung ist und das rezensierte Werk als paradigmatischen Gegenbeweis anführt. In beiden Fällen wird Nichteindeutigkeit rhetorisch erzeugt und genutzt, ohne dass sie im eigentlichen Sinne das Werturteil beeinflusst. Solche Fälle wurden von den Codierenden in der Regel auch holistisch als eindeutig positive Wertungen erkannt und nicht in ihre kontrastiven rhetorischen Bestandteile zerlegt.

[41]

Zweitens können unterschiedlich wertende Aussagen innerhalb einer Rezension durch das unterschiedliche Gewichten dieser Werturteile in ein numerisch eindeutiges 5-Sterne-Urteil überführt werden. Die betreffenden Attribute müssen dafür als minimal und maximal bedeutsam beurteilt werden. Die »Bedeutsamkeit« ist dabei eine subjektive Kategorie, allerdings scheinen inhaltliche und sprachliche Aspekte in der Regel als relevanter eingestuft zu werden als materielle, para- und kontextuelle oder Kriterien der Distribution (d. h. der Zustand eines Buches, Klappentext, formale Genrezuordnung, Titel oder Lieferung). Hierin zeigt sich, dass bei digitalen Rezensionspraktiken ein Konsens darüber zu bestehen scheint, dass im Falle von Büchern weniger die Dienstleistung des Verkäufers oder das physische Produkt zu bewerten sind, sondern die literarische Qualität oder weiter gefasst der Inhalt eines Werkes sowie die Rezeptionserfahrung. Nichteindeutigkeit resultiert in solchen Fällen also aus der ontologischen Mehrfachcodiertheit des Buches als einem materiellen Gegenstand, einem käuflichen Produkt und Kulturgut. In beiden Fällen liegen somit Formen der Nichteindeutigkeit vor, die sich nicht unmittelbar auf die abschließende Beurteilung des Rezensionsobjektes auswirken, weshalb die sprachlich-inhaltliche Ebene keinen Widerspruch zur positiven numerischen Indexierung erzeugt. [42]

5.2 Nichteindeutigkeit in 1- und 2-Sterne-Rezensionen

Auch in numerisch eindeutig negativen Rezensionen (n = 6) treten sprachlich-inhaltliche Nichteindeutigkeiten auf. Rezensionen, die als vollständige »Verrisse« ausschließlich und mehrfach eindeutig negative Wertaussagen vornehmen, wie es das Beispiel (3) zeigt, sind insgesamt in der Unterzahl (n = 2). [43]

(3) [44]

»Der Jugendroman »Tschick« hat mich nicht wirklich überzeugt, er war mir das Geld nicht Wert. Ich kann ihn nicht weiterempfehlen, da mir vieles nicht logisch erscheint und mir die Story nicht gefällt. [...] Zudem gefällt mir die Jugendsprache nicht [...].«

(Zaneta Czamarczan-Marach / Amazon: Rezension zu *Tschick* von Wolfgang Herrndorf, 1 Stern. 08.02.2015. HTML. [online])

Beispiel (3) verdeutlicht, dass mehrfaches eindeutig negatives Werten innerhalb einer Rezension sowohl in der Wiederholung, Variation oder Steigerung einer Aussage (»nicht [...] überzeugt Geld nicht Wert«, »nicht weiterempfehlen«) oder eines Kritikpunktes bestehen kann, zudem aber auch verschiedene Attribute des Werkes (und nicht nur *ein* schwerwiegender disqualifizierender Makel) genannt werden, die negativ gewertet werden (z. B. Missfallen des Aufbaus, der Erzähltechnik und des Stils). Dies spricht dafür, dass der negativen Wertung eine literaturtheoretische Vorstellung davon zugrunde liegt, welche Aspekte bei einer Buchrezension zu berücksichtigen sind. Auch wenn in diesem Fall das Gefühl des »Missfallens« nicht eingehend begründet wird und es somit bei einer klassischen Gefallenswertung⁴¹ bleibt, besteht doch immerhin das Bestreben, die Wertung an unterschiedliche textuelle Charakteristika zurückzubinden. [45]

In der Regel waren die untersuchten Rezensionen mit niedrigen Sternewerten dagegen sprachlich-inhaltlich ausgewogener oder konditional formuliert. So differenziert der / die Rezensent*in in Beispiel (4) sehr klar zwischen der Buchidee / dem narrativen Potenzial und der Umsetzung: [46]

(4) [47]

⁴¹ Zur Unterscheidung von Gefallens- und Anerkennungswertung in der psychologischen Wertungsforschung vgl. van Holt 2013, S. 24.

»[...] Die Story ist außergewöhnlich, leider ist der Schreibstil des Autors sehr schlecht. Die Figuren sind flach und wenig sympathisch, die Sprache soll wohl ›hip‹ sein, ist es aber nicht. Die Idee war gut, es hätte ein tolles Buch sein können, die Umsetzung aber leider banal«

(Lesefreundin / Amazon: Rezension zu *Die Auserwählten – Im Labyrinth: Maze Runner 1* von James Dashner, 2 Sterne. 26.11.2014. HTML. [\[online\]](#))

Ähnlich wie schon im Fall der 5-Sterne-Rezensionen werden die beiden komplementär wertenden Aussagen im Beispielfall (4) (›gute Idee‹ vs. ›schlechte Umsetzung‹) unterschiedlich gewichtet, sodass die abschließende (auch numerische) Einordnung eindeutig negativ ausfällt. Dass dabei der ›Umsetzung‹ des narrativen Potenzials deutlich mehr Gewicht beigemessen wird als der ›Idee‹ an sich, scheint im Rahmen einer an der Leseerfahrung interessierten Rezensions- und Empfehlungspraxis folgerichtig. Es ließen sich aber auch literarische Werke, Gattungen, Genres und Rezensionsformen denken, in denen eine solche Gewichtung anders ausfallen könnte (wie beispielsweise bei Konzept- oder Thesenromanen sowie einer diskursiv ausgerichteten Literaturkritik). Wie bereits in Beispiel (3) werden zudem sehr unterschiedliche Attribute (Figurengestaltung, Sprachstil, Umsetzung insgesamt) herangezogen, um die negative Gesamtwertung differenziert zu begründen. Auf den ersten Blick und auch dem Urteil der Codierenden zufolge, liegen hier keine Nichteindeutigkeiten vor. Legt man allerdings ein strengeres, literaturwissenschaftliches Raster an die Aussagen, ergeben sich ungewollte Nichteindeutigkeiten durch unklare Begriffsverwendung und Vagheit: So bleibt beispielsweise offen, was eine »banal[e Umsetzung]« sein soll. Als konkrete Aspekte dieser »Umsetzung« werden lediglich die Figurengestaltung und der Sprachstil genannt. Eine Leerstelle bleibt, welches erzählerische Vorgehen hier genau als problematisch angesehen wird: Ist der Roman zu schematisch, zu vorhersehbar, überstrukturiert, nicht spannungsreich genug etc. In der Kommunikation über Bücher scheinen, so zeigen es die numerischen Urteile und auch die Codierungen, derartige Nichteindeutigkeiten aber eine untergeordnete Rolle zu spielen. [48]

In einigen 1- und 2-Sterne-Rezensionen entstehen Nichteindeutigkeiten in Bezug auf die Wertung dadurch, dass die Urteile in Relation zu einer bestimmten Rezeptionserwartung (für eine Rezipient*innengruppe) oder zu anderen Konditionen gesetzt werden. In Beispiel (5) wird das negative Urteil dadurch relativiert, dass es auf ein bestimmtes Genre bezogen wird. Charlotte *Links Die letzte Spur* enthält der Rezensent*in »[a]ls Krimi« zu wenig Spannung: [49]

(5) [50]

»[...] Das war es dann auch. Als Krimi liefert mir das Buch zu wenig Spannung. [...] Ich erwartete ein 2-Sterne-Buch und habe es bekommen.«

(Marie / BücherTreff.de: Rezension zu *Die letzte Spur* von Charlotte Link, 2 Sterne. 27.11.2009. HTML. [\[online\]](#))

Andere Rezensionen dieser Art bedingen ihr Urteil anhand der Zielgruppe (»Es ist ja eigentlich ein Kinderbuch.«) oder der Erwartungshaltung, die von einer Buchverfilmung angeregt worden war (»Meine Erwartungen waren nach der tollen Verfilmung wohl etwas zu hoch.«). Das negative numerische Urteil bleibt von der Relativierung allerdings zumeist unberührt. Die Rezensent*innen setzen damit ihre eigene Lektüreerfahrung numerisch absolut, während sie auf der sprachlich-inhaltlichen Ebene differenzierter vorgehen. Während dieses ›Verschwinden‹ der Nichteindeutigkeiten in der numerischen Wertung in den bisher betrachteten Fällen zumeist einer gegenstandsbezogenen oder funktionalen Logik bzw. Gewichtung folgte, ist der Wegfall der *Wertungskondition* in diesen Fällen schwerer zu begründen. Zu vermuten ist, dass eine inhärente Konzeption der Sterne-Symbolik als einer subjektiven Erfahrungseinheit es erlaubt, die einschränkenden Aspekte der eigenen Rezeptionserfahrung zu vernachlässigen. Problematisch ist dies nur insofern, als die numerische Skala solche individuellen Kontexte nicht nur nicht abbildet, sondern diese systematisch hinter der vermeintlichen Objektivität der Zahlen verbirgt. [51]

Gesteigert ist diese Diskrepanz noch in Fällen, in denen die Rezensent*innen mit ihrer numerischen *und* der sprachlich-inhaltlichen Wertung von dem funktionalen Konzept abweichen, das der Mehrheit der Buchrezensionen auf den untersuchten Plattformen zugrunde liegt – primär eine Auseinandersetzung mit bzw. Bewertung des *Buchinhaltes* vorzunehmen: [52]

(6) [53]

»Ich habe das buch bestellt und es ist leider nicht so eingetroffen wie es beschrieben wurde der zustand war schrecklich abgenutzt so das [sic] man keine lust mehr hat es zu lesen! [sic]«

(amira sharif / Amazon: Rezension zu *Deutschland schafft sich ab* von Thilo Sarrazin, 1 Stern. 06.11.2014. HTML. [online])

Neben der Diskrepanz zur mehrheitlichen Rezensionspraxis veranschaulicht Beispiel (6) die Nichteindeutigkeit einer Wertung, die als inhaltliche Leerstelle auftritt. Denn in dieser Rezension kommuniziert nicht ein Subjekt als empfehlende Leser*in, sondern eine enttäuschte Amazon-Kundin, die den physischen Zustand eines Produktes bemängelt. Über den Inhalt des Buches, macht die Verfasser*in keine Aussage – sie stellt also keine Eindeutigkeit in Bezug auf die auf der Plattform eigentlich konventionalisierte Wertungskategorie ‚Buchinhalt‘ her. [54]

Wie bereits im Falle der 5-Sterne-Rezensionen lässt sich auch in Bezug auf die qualitativ analysierten 1- und 2-Sterne-Rezensionen von Amazon und BücherTreff.de zusammenfassend festhalten, dass die sprachlich-inhaltlich kommunizierten Nichteindeutigkeiten häufig ohne Auswirkungen auf die Gesamtwertungen bleiben und nicht immer codiert wurden. Umgekehrt finden sich auf der Aussageebene mehr Nichteindeutigkeiten, als es die eindeutigen numerischen Werte vermuten lassen. Auch hierunter fanden sich Fälle, die aufgrund individueller oder kollektiver Gewichtung bei der Gesamtbewertung nicht berücksichtigt wurden, aber auch nicht-intendierte begriffliche Inkonsistenzen sowie konditionale Einschränkungen von Wertungen. [55]

5.3 Werksspezifische Nichteindeutigkeit

In einigen der untersuchten Rezensionen scheinen weder rhetorische noch soziale Funktionen, auch nicht konditionale Einschränkungen, begriffliche Ungenauigkeiten oder komplementäre textuelle Befunde die Nichteindeutigkeit der Wertung zu motivieren. Stattdessen zielt das Buch selbst darauf, Wertekonflikte in der Rezeptionserfahrung auszulösen. Mehr als einmal bezogen sich solche Rezensionen auf die satirische Dystopie von Timur Vermes *Er ist wieder da*. Der Unterhaltungsroman ergründet satirisch anhand eines Zeitreis Szenarios, welche Auswirkungen die Wiederkehr Adolf Hitlers in der Gegenwart der 2000er-Jahre zeitigen könnte. Die Antwort bleibt warnend-ambivalent und humoristisch eingebettet, wenn sie Hitlers Karriere zwischen Comedy-Star und populistischem Brandstifter 2.0 verortet. Der Roman bedient sich damit in Deutschland gesellschaftlich und erinnerungskulturell sensibler und teilweise auch tabuisierter Themen und verarbeitet diese in einem Werk der Unterhaltungsliteratur häufig mit dem Mittel des parodistischen Humors. Sowohl die Gefahr einer Verharmlosung des Nationalsozialismus als auch die humoristische Relativierung aktueller populistischer Politik bilden verbreitete Kritikpunkte am Roman. Die Frage der Angemessenheit eines solchen Vorgehens stellt sich dabei nicht nur auf einer Ebene des gesellschaftlichen Diskurses, sondern – so belegen es die Rezensionen – in vielen individuellen Rezeptionserfahrungen: [56]

(7) [57]

»Aber schlussendlich habe ich es nicht bereut. Es war interessant und ja ich muss zugeben, auch auf eine ganz komische (und beängstigende) Art witzig.«

(Nicole / Amazon: Rezension zu *Er ist wieder da* von Timur Vermes, 5 Sterne. 08.04.2013. HTML. [online])

Eine milde Form des Wertekonflikts wird hier über den Einschub »(und beängstigende) Art« sowie über die Gewichtung vorangegangener Einwände im Gesamturteil (»schlussendlich«) markiert. Die darin ausgedrückte Nichteindeutigkeit kann hier also trotzdem zu einem eindeutig positiven numerischen Urteil führen. Der Rezensentin / dem Rezensenten aus Beispiel (8) ist das Buch allerdings nur einen Stern wert. Dabei liest sich die Begründung sehr ähnlich wie jene aus Beispiel (7): [58]

(8) [59]

»Das ganze Buch ist sehr nett geschrieben. [...] Am Ende des Buchs hat man diesen verschrobenen Protagonisten irgendwie lieb gewonnen und wünscht ihm, dass er seinen Weg gehen wird. Und genau darin liegt für mich die Gefahr!«

(harald roban / Amazon: Rezension zu *Er ist wieder da* von Timur Vermes, 1 Stern. 21.06.2015. HTML. [online])

Die hier vorliegende Nichteindeutigkeit entsteht unmittelbar aus dem intraindividuellen Konfligieren zweier unterschiedlicher Werte, nämlich einerseits einem moralischen und andererseits einem Unterhaltungswert. Nicht zuletzt zeigt sich an diesen Beispielen, dass es offensichtlich Bücher gibt, die aufgrund solcher inhärenten Konfliktpotenziale zwischen verschiedenen Werten auch Nichteindeutigkeiten bei ihrer Rezeption provozieren. [60]

5.4 Nichteindeutigkeit in 3-Sterne-Rezensionen

Die überwiegende Mehrzahl der 3-Sterne-Rezensionen (insgesamt n = 21) fällt auf der Textebene tatsächlich sowohl positive als auch negative Wertungen über das gelesene Buch. Die Werke werden also als »mittelmäßig« eingestuft, wie es in einer Rezension auch explizit heißt. Sowohl für die Interpretation der numerischen Einordnung als auch für die Codierungen lässt sich somit also eine gewisse Validität nachweisen: Drei vergebene Sterne und auch das Codieren als nichteindeutige Rezension durch die Codierenden waren gute Hinweise auf tatsächliche Nichteindeutigkeiten der Wertungen in den Rezensionen. Die häufigste Ursache für nichteindeutige Wertungen oder polare Wertungen in Bezug auf ein Werk liegt in komplementären Textbefunden (Beispiel 9) oder Wertungskriterien (10). Die textuelle Manifestation eines bestimmten Wertungskriteriums (wie Immersion) wird beispielsweise am Anfang des Buches als hoch, später als gering empfunden (9): [61]

(9) [62]

»[...] Der Anfang ist detailreich beschrieben, gemeinsam mit Chris, dem Protagonisten, bewegt sich der Leser über das Partygelände, als Betrachter. Schon bald begegnen wir Jenny, und ab da verflog meine anfängliche Begeisterung von Seite zu Seite. Denn anstatt irgendwann richtig ins Geschehen einzutauchen, bleibe ich in einer beobachtenden Position. [...]«

(Knü / BücherTreff.de: Rezension zu *Tödliche Missverständnisse* von Philip Pullman, 3 Sterne)

Im nachfolgenden Beispiel wird nicht ein Aspekt im Textverlauf betrachtet, sondern es werden verschiedene Aspekte des Buches (Figuren, Handlung etc.) als unterschiedlich »gelingen« empfunden. »Gelingen« kann sich dabei auf unterschiedliche Wertungskriterien für verschiedene formale und inhaltliche Dimensionen beziehen, wie beispielsweise literarästhetisch auf »grandiose Beschreibungen« und wirkungsorientiert auf langweilige Ausschweifungen.⁴² [63]

(10) [64]

⁴² Die Unterscheidung orientiert sich an der Typologie axiologischer Werte nach von Heydebrand / Winko 1996, S. 114–115.

»[...] Die Beschreibungen der verschiedenen Orte und Gestalten ist [sic] grandios und sucht sicherlich seinesgleichen [sic]. Im Grunde sind diese Beschreibungen ein ganz dicker Pluspunkt und wohl 5 Sterne wert, wären da nicht die ständigen Wiederholung [sic] bereits angesprochener Themen und die extremen Ausschweifungen. [...]«

(Dave2311 / BücherTreff.de: Rezension zu *Berge des Wahnsinns* von H. P. Lovecraft, 3 Sterne. 13.05.2016. HTML. [\[online\]](#))

In den beiden letzten Beispielen werden die konfligierenden Wertungen nicht, wie im Falle der 1- und der 5-Sterne-Rezensionen, am Ende zu einer eindeutigen Wertung gewichtet. Sie bleiben stattdessen gleichwertig nebeneinander bestehen und führen zu einem sprachlich ambivalenten und numerisch »mittelmäßigen« Urteil.

[65]

6. Nichteindeutigkeit als intraindividueller Wertungsdissens – ein Fazit

Zusammenfassend lässt sich in Bezug auf die Eindeutigkeit von Wertungspraktiken in Online-Rezensionen auf Amazon und BücherTreff.de festhalten, dass eine große Mehrzahl der wertenden Aussagen eindeutig positiv ist (ca. 80 % im Durchschnitt beider Plattformen). Auf Rezensionsebene bleibt diese hohe Positiv-Rate erhalten, allerdings verändert sich das Verhältnis von nichteindeutigen Rezensionen zu eindeutig negativen Rezensionen zugunsten der nichteindeutigen (von ca. 10,2 % auf Aussageebene zu 21,2 % auf Rezensionsebene im Durchschnitt beider Plattformen). Dies lässt sich dadurch erklären, dass negative Aussagen in Rezensionen selten unwidersprochen, sondern zu einem Großteil in Kombination mit positiven Aussagen auftreten. Im Hinblick auf numerische Wertungspraktiken konnte auf Amazon tendenziell eine Übereinstimmung zwischen dem Grad der Eindeutigkeit im Rezensionstext und der Anzahl der vergebenen Sterne festgestellt werden. Auf BücherTreff.de überwiegen eindeutig positive textuelle Wertungshandlungen unabhängig von den vergebenen Sternen. Die Befunde, die also eine leichte Plattformspezifität in Bezug auf die Nichteindeutigkeit von Wertungen belegen konnten, stimmen mit bereits bekannten Forschungsergebnissen überein, die Online-Rezensionen als tendenziell positive Wertungspraxis charakterisieren und die das Praxisfeld als multifaktoriell beeinflusst beschreiben. Die Sprachhandlungen werden also nicht nur durch die Plattformkonvention determiniert, sondern vom Zusammenspiel verschiedener Faktoren (wie beispielsweise individuellen Kompetenzen, dem rezensierten Artefakt etc.). In einem nächsten Schritt wurden daher weitere Einflussvariablen auf die Eindeutigkeit von Online-Rezensionen untersucht.

[66]

Die Analyse der Eindeutigkeit von Bewertungen erwies sich zudem als durch die Messinstrumente beeinflussbar. Beim intuitiven Beurteilen der Online-Rezensionen durch menschliche Codierende wurden die Bewertungen als eindeutiger wahrgenommen, als sie sich auf Basis rein automatisierter (sentiment-analytischer) und qualitativer Analyseverfahren darstellten. Dies spricht dafür, dass Nichteindeutigkeiten von Bewertungen in Online-Rezensionen kontextabhängig aufgelöst werden können und sie insbesondere dann, wenn sie in schwacher Form auftreten, nicht unbedingt von Rezipient*innen als solche wahrgenommen werden. Perspektivisch wären daher weitere Betrachtungen der Ergebnisse der Sentiment-Analyse unter Berücksichtigung des Kontextes und der verwendeten Lemmata wünschenswert.

[67]

Nichteindeutigkeit hat in Online-Rezensionen verschiedene Funktionen, die von rhetorischen über reflektierend-gewichtende bis hin zu inhaltlich-differenzierenden reichen. Aber auch nichtintendierte Nichteindeutigkeiten sind anzunehmen, die beispielsweise wegen der Mehrfachkodiertheit des Buches als Konsumprodukt und Kulturgut oder aufgrund begrifflicher Inkonsistenzen entstehen. In einigen der untersuchten Rezensionen gehen Nichteindeutigkeiten aus intraindividuellen Wertungskonflikten hervor, die das rezensierte Werk mitunter gezielt durch seine textuelle Performanz konfligierender Werteordnungen erzeugt. In Anlehnung an die Vorschläge von von Heydebrand und Winko, die als Ursache

[68]

für den interindividuellen Dissens zwischen literarischen Werturteilen die vier Problematiken der unklaren / konfligierenden Textbeschreibung, axiologischen Werte, Wertungskriterien und Gewichtung der Werte sehen,⁴³ lassen sich ggf. auch die intraindividuellen Ursachen für Nichteindeutigkeiten systematisieren. So fanden sich in den Beispielen Belege für Nichteindeutigkeiten sowohl aufgrund unklarer Textbezüge ebenso wie aufgrund sich widersprechender Werte (vgl. Unterhaltung vs. Moral). Auch verschiedene Wertungskriterien führten teilweise zu ambivalenten Urteilen in Bezug auf ein und dasselbe Werk (wie Landschaftsbeschreibungen vs. Ausschweifungen). Umgekehrt konnte eine einseitige Gewichtung der Werte oder -kriterien Nichteindeutigkeiten in ein abschließend eindeutiges Urteil überführen. Die Beantwortung der Fragen, inwiefern sich diese Auffächerung nach wertungstheoretischen Ursachen für Nichteindeutigkeiten auch in anderen Rezensionsformaten wiederfinden lässt, ob sie über Plattform- und Genre-Grenzen hinaus Anwendung findet und welche Rolle die oben genannten Funktionen von Nichteindeutigkeiten von Wertungen in Online-Rezensionen dabei spielen, bleibt ein Desiderat der zukünftigen Forschung. Die bisherigen Untersuchungsergebnisse liefern dabei bereits wichtige Hinweise darauf, dass sich das pauschale Vorurteil einer Vorherrschaft von ›Lob und Torheit‹ im Bereich digitaler Lai*innenrezensionen im Sinne unterkomplexer bzw. ausschließlich (und damit eindeutig) positiv wertender Rezensionen so nicht halten lässt. Denn auch wenn sich viele eindeutig positive Rezensionen auf Amazon und BücherTreff.de finden ließen, so zeigt die Untersuchung doch auch, dass die Eindeutigkeit literarischer Wertung offensichtlich von weitaus mehr Faktoren abhängt als dem Äußerungsort oder der angenommenen Wertungskompetenz einer bestimmten Personengruppe.

⁴³ von Heydebrand / Winko 1996, S. 105–110.

Bibliografische Angaben

aboutbooks GmbH (Hg.): LovelyBooks. Letzter Zugriff: 12.06.2024. HTML. [\[online\]](#)

Daniel Allington: »Power to the Reader« or »Degradation of Literary Taste«? Professional Critics and Amazon Customers as Reviewers of »The Inheritance of Loss«. In: Language and Literature. International Journal of Stylistics 25 (2016), H. 3, S. 254–78. DOI: 10.1177/0963947016652789

Thomas Anz: Literaturkritik als (Neben-)Beruf. Informationen und Anleitungen zu Praxis. In: Thomas Anz / Rainer Baasner (Hg.): Literaturkritik. Geschichte, Theorie, Praxis (= Beck'sche Reihe, 1588). München 2004, S. 220–236. [\[Nachweis im GVK\]](#)

Thomas Anz: Kontinuitäten und Veränderungen der Literaturkritik in Zeiten des Internets. Fünf Thesen und einige Bedenken. In: Renate Giacomuzzi / Stefan Neuhaus / Christiane Zintzen (Hg.): Digitale Literaturvermittlung. Praxis, Forschung und Archivierung. Innsbruck u. a. 2010, S. 48–59. [\[Nachweis im GVK\]](#)

Andrea Bachmann-Stein: Zur Praxis des Bewertens in Laienrezensionen. In: Heinrich Kaulen / Christina Gansel (Hg.): Literaturkritik heute. Tendenzen – Traditionen – Vermittlung. Göttingen 2015, S. 77–91. [\[Nachweis im GVK\]](#)

Moritz Baßler: Der neue Midcult. Vom Wandel populärer Leseschichten als Herausforderung der Kritik. In: Pop-Zeitschrift 10 (2021), H. 1, S. 132–149. DOI: 10.14361/pop-2021-100122

Moritz Baßler / Heinz J. Drügh: Gegenwertsästhetik. Göttingen, Konstanz 2021.

Peter Boot / Marijn Koolen: Captivating, Splendid or Instructive? Assessing the Impact of Reading in Online Book Reviews. In: Scientific Study of Literature 10 (2020), H. 1, S. 66–93. [\[Nachweis im GVK\]](#)

Börsenblatt: Medien und Macher. Treffen im Land der Bücher. 27.01.2016. HTML. [\[online\]](#)

Axel Bruns: Blogs, Wikipedia, Second Life, and Beyond: From Production to Produsage (= Digital Formations, 45). New York u. a. 2008. [\[Nachweis im GVK\]](#)

BücherTreff.de. Das volle Bücherleben. Letzter Zugriff: 10.04.2024. HTML. [\[online\]](#)

Judith A. Chevalier / Dina Mayzlin: The Effect of Word of Mouth on Sales. Online Book Reviews. In: Journal of Marketing Research 43 (2006), H. 3, S. 345–354. DOI: 10.1509/jmkr.43.3.345

Stefan Dimitrov / Faiyaz Zamil / Andrew Piper / Derek Ruths: Goodreads versus Amazon. The Effect of Decoupling Book Reviewing and Book Selling. In: Proceedings of the Ninth International AAAI Conference on Web and Social Media. (International AAAI Conference on Web and Social Media 9, Oxford, UK, 26.–29.05.2015) Oxford, UK 2015, S. 602–605. PDF. [\[online\]](#)

Beth Driscoll / DeNel Rehberg Sedo: Faraway, So Close. Seeing the Intimacy in Goodreads Reviews. In: Qualitative Inquiry 25 (2018), H. 3, S. 1–12. DOI: 10.1177/1077800418801375

Jakob Fehle / Thomas Schmidt / Christian Wolff: Lexicon-based Sentiment Analysis in German: Systematic Evaluation of Resources and Preprocessing Techniques. In: Kilian Evang / Laura Kallmeyer / Rainer Osswald / Jakob Waszczuk / Torsten Zesch (Hg.): Proceedings of the 17th Conference on Natural Language Processing. (KONVENS 2021, Düsseldorf, 06.–09.09.2021) Düsseldorf 2021, S. 86–103. PDF. [\[online\]](#)

Renate Giacomuzzi: Die »verkehrte« Welt der Literaturblogs. Zur Funktion und Position »freier« Literatur- und Leserblogs im Feld der Literaturkritik. In: Doris Moser / Claudia Dürr (Hg.): Über Bücher reden. Literaturrezeption in Lesegemeinschaften (= Digilit, 3). Göttingen 2021, S. 183–197. PDF. DOI: 10.14220/9783737013239.183

Barney G. Glaser / Anselm L. Strauss: Grounded theory. Strategien qualitativer Forschung. 3. Auflage. Bern 2010. [\[Nachweis im GVK\]](#)

Guido Graf / Ralf Knackstedt / Kristina Petzold (Hg.): Rezension – Online-Rezensionen und Kulturelle Bildung (= Digital Humanities, 2). Bielefeld 2021. DOI: 10.1515/9783839454435

Renate von Heydebrand / Simone Winko: Einführung in die Wertung von Literatur. Systematik – Geschichte – Legitimation (= Literaturwissenschaft, 1953). Paderborn u. a. 1996. 216011728

Nadine van Holt: Psychologische Werttheorien. In: Gabriele Rippl / Simone Winko (Hg.): Handbuch Kanon und Wertung. Theorien, Instanzen, Geschichte. Stuttgart 2013, S. 19–24. [\[Nachweis im GVK\]](#)

Rebecca E. Johnson: The New Gatekeepers. How Blogs Subverted Mainstream Book Reviews. Dissertation, Virginia Commonwealth University. 2016. Richmond, US-VA 2016. PDF. [\[online\]](#)

Holger Kellermann / Gabriele Mehling: Laienrezensionen auf amazon.de im Spannungsfeld zwischen Alltagskommunikation und professioneller Literaturkritik. In: Andrea Bartl / Markus Behmer (Hg.): Die Rezension. Aktuelle Tendenzen der Literaturkritik (= Konnex, 22). Würzburg 2017, S. 173–202. [\[Nachweis im GVK\]](#)

Wendy Kerstan: Der Einfluss von Literaturkritik auf den Absatz von Publikumsbüchern. Marburg 2006. [\[Nachweis im GVK\]](#)

Regina Kirchmeier: Bloggen und Kooperationen: Aus der Perspektive von Mikro-Influencern. In: Annika Schach / Timo Lommatzsch: Influencer Relations. Marketing und PR mit digitalen Meinungsführern. Wiesbaden 2018, S. 303–314. [\[Nachweis im GVK\]](#)

Raphaela Knipp: Literaturbezogene Praktiken. Überlegungen zu einer praxeologischen Rezeptionsforschung. In: Navigationen. Zeitschrift für Medien- und Kulturwissenschaften 17 (2017), H. 1, S. 95–116. PDF. DOI: 10.25969/mediarep/1736

Marijn Koolen / Peter Boot / Joris J. van Zundert: Online Book Reviews and the Computational Modelling of Reading Impact. In: Folger Karsdorp / Barbara McGillibray / Adina Nerghes / Melvin Wevers (Hg.): Proceedings of the Workshop on Computational Humanities Research (CHR 2020). (CHR 2020, Amsterdam, 18.–20.11.2020) Amsterdam 2020, S. 149–169. PDF. [\[online\]](#)

Kristin Kutzner / Kristina Petzold / Ralf Knackstedt: Characterising Social Reading Platforms. A Taxonomy-Based Approach to Structure the Field. In: Wi 2019. Wirtschaftsinformatik 2019 Proceedings. (14th International Conference on Wirtschaftsinformatik, Siegen 24.–27.02.2019), S. 676–690. PDF. [\[online\]](#)

Kristin Kutzner / Anna Moskvina / Kristina Petzold / Claudia Roßkopf / Ulrich Heid / Ralf Knackstedt: Reviews of Cultural Artefacts: Towards a Schema for their Annotation. In: Sandra Kübler / Heike Zinsmeister (Hg.): Proceedings of the Workshop on Annotation in Digital Humanities co-located with ESLLI 2018. (annDH 2018, Sofia, 6.–10.08.2018) 2018, S. 17–23. PDF. [\[online\]](#)

Piroska Lendvai / Simone Rebora / Moniek Kuipers: Identification of Reading Absorption in User-Generated Book Reviews. In: Proceedings of the 15th Conference on Natural Language Processing. (KONVENS 2019, Erlangen / Nürnberg 9.–11.10.2019) Erlangen 2019, S. 271–272. PDF. [\[online\]](#)

Niklas Luhmann: Die Kunst der Gesellschaft. Frankfurt am Main 1997.

Thorben Mämecke: Das quantifizierte Selbst. Zur Genealogie des Self-Trackings (= Digitale Gesellschaft, 34). Bielefeld 2021. DOI: 10.1515/9783839456033

Steffen Mau: Das metrische Wir. Über die Quantifizierung des Sozialen. Berlin 2017. [\[Nachweis im GVK\]](#)

Ronan McDonald: The Death of the Critic. London u. a. 2007. [\[Nachweis im GVK\]](#)

Gabriele Mehling / Axel Kellermann / Holger Kellermann / Martin Rehfeldt: Leserrezensionen auf amazon.de. Eine teilautomatisierte inhaltsanalytische Studie. Bamberg 2018. [\[Nachweis im GVK\]](#)

Doris Moser / Claudia Dürr (Hg.): Über Bücher reden. Literaturrezeption in Lesegemeinschaften (= Digilit, 3). Göttingen 2021. DOI: [10.14220/9783737013239](#)

Anna Moskvina / Kristina Petzold: Online Reviews as Communal Formations. A Corpus Linguistic Approach. In: Journal of Cultural Analytics. Special Issue [im Erscheinen].

Stefan Neuhaus: »Leeres, auf Intellektualität zielendes Abrakadabra«. Veränderungen von Literaturkritik und Literaturrezeption im 21. Jahrhundert. In: Heinrich Kaulen / Christina Gansel (Hg.): Literaturkritik heute. Tendenzen – Traditionen – Vermittlung. Göttingen 2015, S. 43–57. [\[Nachweis im GVK\]](#)

Katharina Evelin Perschak: »Ich lese mit.« Kommunikation und Interaktion in Online-Lesegruppen. In: Doris Moser / Claudia Dürr (Hg.): Über Bücher reden. Literaturrezeption in Lesegemeinschaften (= Digilit, 3). Göttingen 2021, S. 103–129. DOI: [10.14220/9783737013239.103](#)

Kristina Petzold: Buchblogs zwischen Passion und Profession. Zur Diskursivierung digitaler literaturbezogener Anschlusskommunikation als Arbeit. In: digilit 5 (2025), V&R unipress. DOI: [10.14220/9783737016667](#).

Stephan Porombka: Weg von der Substanz. Hin zu den Substanzen. In: Matthias Beilein / Claudia Stockinger / Simone Winko (Hg.): Kanon, Wertung und Vermittlung. H. Berlin u. a. 2011, S. 293–303. DOI: [10.1515/9783110259964.293](#)

Simone Rebora / Peter Boot / Federico Pianzola / Brigitte Gasser / J. Berenike Herrmann / Maria Kraxenberger / Moniek M. Kuijpers / Gerhard Lauer / Piroška Lendvai / Thomas C. Messerli / Pasqualina Sorrentino: Digital Humanities and Digital Social Reading. In: Digital Scholarship in the Humanities 36 (2021), Supplement 2, S. ii230–ii250. DOI: [10.1093/llc/fqab020](#)

Simone Rebora / Piriska Lendvai / Moniek Kuijpers: Reader Experience Labeling Automatized. Text Similarity Classification of User-Generated Book Reviews. In: EADH2018 Book of Abstracts (2018). PDF. [\[online\]](#)

Martin Rehfeldt: »Ganz große, poetische Literatur – Lesebefehl!« Unterschiede und Gemeinsamkeiten von Amazon-Rezensionen zu U- und E-Literatur. In: Sebastian Böck / Julian Ingelmann / Kai Matuszkiewicz / Friederike Schruhl (Hg.): Lesen X.0. Rezeptionsprozesse in der digitalen Gegenwart. Göttingen 2017, S. 235–250. PDF / EPUB. DOI: [10.14220/9783737007450.235](#)

Vanessa-Isabelle Reinwand-Weiss / Claudia Roßkopf: Erkenntnisse aus bildungstheoretischer Sicht. In: Guido Graf / Ralf Knackstedt / Kristina Petzold (Hg.): Rezensiv – Online-Rezensionen und Kulturelle Bildung (= Digital Humanities, 2). Bielefeld 2021, S. 79–109. DOI: [10.1515/9783839454435](#)

Robert Remus / Uwe Quasthoff / Gerhard Heyer: SentiWS. A Publicly Available German-language Resource for Sentiment Analysis. In: Nicoletta Calzolari / Khalid Choukri / Bente Maegaard / Joseph Mariani / Jan Odijk / Stelios Piperidis / Mike Rosner / Daniel Tapias (Hg.): Proceedings of the Seventh International Conference on Language Resources and Evaluation. (LREC'10, Valletta, MT Mai 2010) Valletta, MT 2010. PDF. [\[online\]](#)

Christine Römer / Brigitte Matzke: Lexikologie des Deutschen. Eine Einführung. 2. Auflage. Tübingen 2005. 475143493

Siegfried J. Schmidt: Grundriss der empirischen Literaturwissenschaft. Frankfurt am Main 1991.

Barbara Herrnstein Smith: What Was »Close Reading«? A Century of Method in Literary Studies. In: The Minnesota Review (2016), H. 87, S. 57–75. DOI: [10.1215/00265667-3630844](#)

Johannes Spengler: Lesen als Wettbewerb. Selbstvermessung und Challenges von BuchbloggerInnen. In: Tagungsband des Nachwuchstages Poetik der Quantität der Bamberger Graduiertenschule für Literatur, Kultur und Medien. [im Erscheinen]

Bob Stein: A Taxonomy of Social Reading: A Proposal. In: The Insitute for the Future of the Book. 2012. HTML. [\[online\]](#)

Stephan Stein: Laienliteraturkritik. Charakteristika und Funktionen von Laienrezensionen im Literaturbetrieb. In: Heinrich Kaulen / Christina Gansel (Hg.): Literaturkritik heute. Tendenzen – Traditionen – Vermittlung. Göttingen 2015, S. 58–76. [\[Nachweis im GVK\]](#)

Anselm Strauss / Juliet Corbin: Grounded Theory. Grundlagen qualitativer Sozialforschung. Weinheim 1996. [\[Nachweis im GVK\]](#)

Kim M. Sutton / Ina Paulfeuerborn: The Influence of Book Blogs on the Buying Decisions of German Readers. In: Logos. Journal of the World Book Community 28 (2017), H. 1, S. 45–52. DOI: [10.1163/1878-4712-11112124](#)

Étienne Wenger: Communities of Practice. Learning, Meaning and Identity. Cambridge, UK 1998. DOI: [10.1017/CBO9780511803932](#)

Abbildungsverzeichnis

Abb. 1: Ausschnitt aus dem Kategoriensystem / Tagset zur Codierung rezensiver Texte des Hildesheimer Forschungsprojektes Rez@Kultur. [Aus: Graf et al. 2021, S. 401]

Abb. 2: Anteile als nichteindeutig und eindeutig codierter Textabschnitte an den mit diesen Codes versehenen Rezensionen auf Amazon und BücherTreff.de in absoluten und relativen Zahlen.

Abb. 3: Anteile nichteindeutig und eindeutig wertender Rezensionen innerhalb der Gesamtmenge an mit diesen Codes versehenen Rezensionen auf Amazon und BücherTreff.de in absoluten und relativen Zahlen.

Abb. 4: Anteile nichteindeutig und eindeutig wertender Rezensionen verteilt nach der Anzahl der Sterne auf Amazon in absoluten und relativen Zahlen.

Abb. 5: Anteile nichteindeutig und eindeutig wertender Rezensionen verteilt nach der Anzahl der Sterne auf BücherTreff.de in absoluten und relativen Zahlen.

Abb. 6: Anteile von positiven, negativen und nichteindeutigen Rezensionen auf Basis der Sentiment-Analyse auf Amazon und BücherTreff.de.