The Impact of Review Copies on German Online Book Reviews from LovelyBooks

Introduction

Book reviews are of high importance for all stakeholders: For review / book readers, they constitute an important tool to obtain information about books. For publishers and authors, reviews are relevant because they are deemed more credible than traditional marketing strategies. Thus, cooperations between these parties are well established.

An example cooperation bases on incentivization. The most common form consists in giving away free review copies in return for a review. In the last 10 years, literary criticism has experienced a shift: due to social platforms such as LovelyBooks or Goodreads, amateur reviewers become more visible. With respect to the rise of influencers in the book world, cooperations between publishers/authors and book bloggers/bookfluencers are common nowadays.

These book reviewers tend to disclose the cooperations by expressing their gratitude, sometimes by additionally emphasizing that their opinion has not been influenced:

> "Vielen Dank an den Verlag für das Rezensionsexemplar, dies nimmt jedoch keinen Einfluss auf meine Meinung!"

Research Question

Aim
Questioning book reviewers' self-stated objectivity.

Research Question
Do incentivized book reviews show signs of influence if the reviewer received a free book copy?

Background

Book Reviews
Online book reviews, a form of electronic Word of Mouth (eWoM), have become crucial as consumers increasingly rely on peer opinions over traditional marketing. These reviews are particularly valuable in navigating the overwhelming variety of products, including the vast number of about 67,000 newly published books (in Germany 2023).[1]
The increase in reviewing practices influences the whole book world: the rise of amateur literary critics on platforms such as Instagram, TikTok or Goodreads has transformed reading practices into a social, digital exchange, blurring the lines between readers, authors, and critics.
Amateur critics mainly publish book reviews out of these motives: platform assistance, venting extreme feelings, concern for other consumers, self-portrayal, social benefits, economic incentives, helping the author and/or publisher.[2]
While there are multiple platforms to publish book reviews, the currently most popular channels are social cataloging sites such as Goodreads and LovelyBooks. LovelyBooks is the largest provider in the German-speaking area and hosts over 500,000 readers. There are many functionalities, next to sharing and browsing book reviews: for example, users can create individual bookshelves or participate in a reading round.

Incentivization
With respect to the impact of book reviews, there are strategies aimed at fostering review publications. One such exemplary strategy is incentivization, defined as: "The act or process of providing incentives to make something more attractive."[3]
Incentives can have a monetary (e.g. receiving cash) or non-monetary (e.g. receiving a free product) value, a material or immaterial form.
While incentivization is attractive for product owners (e.g. publishers), it can have negative effects on the review reader because most customers read reviews „under the assumption that—unless stated otherwise—content is inherently impartial". To protect review readers from misleading reviews, regulative requirements to disclose sponsorship have been adopted.
In the book world, a popular form of incentivization consists in giving away free book copies to critics in exchange for a review. This thesis focuses on the cooperation between publishers/authors and amateur critics on LovelyBooks. Importantly, review copies constitute a form of non-monetary incentive. In 2019 and out of a legal perspective, this cooperation is a form of advertising. As such, book reviewers are obliged to disclose the fact of having received a free copy. According to the German media state authorities, the following content types are subject to mandatory labeling:
„Beiträge über Produkte, Dienstleistungen, Marken, Unternehmen, Regionen, Events, Reisen, die kostenlos in Anspruch genommen oder erhalten wurden, deren Veröffentlichung aber an Vereinbarungen/Bedingungen geknüpft ist."[4]

Data

Dataset: LoBo Corpus[5]

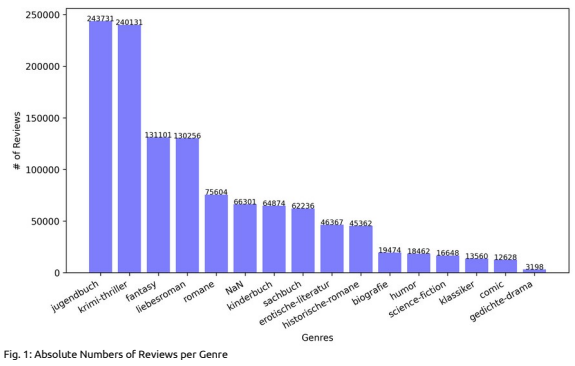

Fig. 1: Absolute Numbers of Reviews per Genre

Metric	Amount
# of reviews	1,327,457
# of genres	16
# of reviewers	54,037
# of reviewed books	169,759
Most reviews per book	1,243
Most reviews per reviewer	55,033

Tab. 1: Descriptive Overview over the LoBo Corpus

Hypotheses

This thesis aims at (partly) replicating the study by Kim et al., 2019.[6] The following hypothesis are transferable to the domain of book reviews:

	Hypothesis	Reasoning
H1	IRs are more positive than NIRs.	Norm of reciprocity
H2	IRs are more complex than NIRs.	Self-fulfilling prophecy
H3	IRs are more elaborate than NIRs.	Self-fulfilling prophecy
H4	IRs are less extreme than NIRs.	Reviewer motivations
H5	IRs are more objective than NIRs.	Reviewer motivations

IR Incentivized review
NIR Non-incentivized review

Explanation

Norm of reciprocity
Refers to Gouldner (1960): „a generalized moral norm of reciprocity which defines certain actions and obligations as repayments for benefits received".[7] Reviewers feel obliged after having received a free copy, which is expressed by positively stating the consumption experience.

Self-fulfilling prophecy
Book reviewers disclose the sponsorship by expressing their gratitude, which reflects a certain consciousness. Thus, the disclosure works as a self-fulfilling prophecy and urges the reviewers to write extraordinarily honest reviews. This could lead to more elaborate and more complex reviews.

Reviewer motivations
NIRs are often written by consumers with a very positive or very negative experience. Thus, IRs are written for other additional purposes and supposed to be less extreme (in their ratings) than NIRs (H4). Equally, NIRs are often published shortly after the experience and as such rather subjective. Contrarily, IRs could be more objective on word level (H5).

Sub-Corpus Building

1. Detection of IRs
A review is incentivized if it contains a disclosure statement with more than 90% accuracy: Rezensionsexemplar, Reziexemplar, Rezi-Exemplar, Freiexemplar, NetGalley, Leseexemplar, Lese-Exemplar

2. Detection of NIRs
A review is non-incentivized if it does not contain a statement from 1.

3. Genre Selection

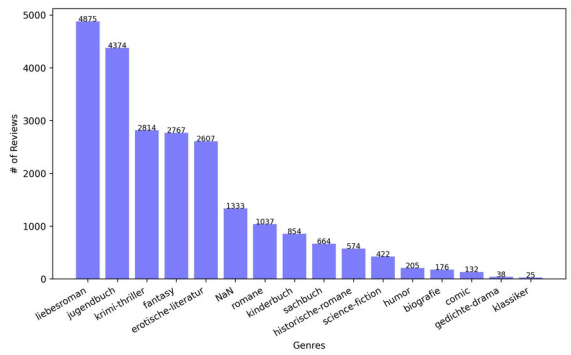

Fig. 2: Absolute Numbers of IRs per Genre

Selected genre: romance

4. Review Type Analysis
There are three review types in the corpus:
1. "Kurzmeinung" (< 140 characters)
2. "Kurzmeinung" & "Rezension"
3. „Rezension"

Due to their length, "Kurzmeinungen" do not convey much meaning and are excluded from the analysis; only type 2 and type 3 reviews are considered.

5. Summary
Sample 1: 4872 IRs (all IRs)
Sample 2: 4872 NIRs (randomized sample)

Preprocessing

1. Unification of quotation marks
2. Deletion of emoticons and special characters
3. Deletion of review type indicator

Hypothesis Testing

Hypothesis	Concept	Formalisation
H1	Positivity	Valence / polarity
H2	Complexity	Average word length
H3	Elaborateness	Review length
H4	Extremeness	Star ratings
H5	Objectivity	Share of non-emotional words

Tab. 2: Formalisation of Hypothesis Concepts

Hypothesis	Concept	Operationalisation
H1	Positivity	AAP (sentiArt[8])*
H2	Complexity	Number of characters**
H3	Elaborateness	Number of words**
H4	Extremeness	Less 1 and 5 star reviews
H5	Objectivity	LIWC-22

Tab. 3: Operationalisation of Hypothesis Concepts

* Only content words are considered
** Only words with a word-related POS-tag are considered (e.g. punctuation is excluded)

Results

Hypothesis	Mean	
	IR	NIR
H1: IRs are more positive than NIRs	0.2612	**0.2775**
H2: IRs are more complex than NIRs	5.2486	**5.2886**
H3: IRs are more elaborate than NIRs	**475.8192**	327.6587
H4: IRs are less extreme than NIRs*	5	**4**
H5: IRs are more objective than NIRs	**91.7514**	91.3303

* Median scores are reported instead because it is an ordinal variable.
Tab. 4: Mean and Median Scores of IRs and NIRs

The mean/median tendency only supports H3 and H4. In the other cases, the hypotheses are rejected and the contrary hypotheses are tested:

H1b NIRs are more positive than IRs.
H2b NIRs are more complex than IRs.
H4b NIRs are less extreme than IRs.

Hypotheses Test Results

Hypothesis	Empirical Value	Probability p	Effect Size
H1b	U = 11546670.0	< 0.05	d = 0.047
H2b	U = 11591870.5	< 0.05	d = 0.04
H3	U = 16043714.0	< 0.001	d = 0.64
H4b	X^2 = 24.9685	< 0.001	-
H5	U = 11868791.0	> 0.05	-

Tab. 5: Hypothesis Test Results

Review Extremity
As the association of star ratings is significant, the relative frequencies of 1- and 5-star ratings are compared:

Sample	Star Rating	
	★	★★★★★
IR	**0.965%**	52.648%
NIR	1.375%	**47.989%**

Tab. 6: Relative Frequencies of Extreme Star Ratings

Interpretation
Overall, four of the original five hypotheses are rejected, among them three due to the wrong tendency as indicated by mean/median scores, one because of lacking statistical significance. H3 is confirmed. The three counter hypotheses are all statistically significant and thus confirmed. While H4b is statistically confirmed, the relative frequencies do not support the hypothesis. Thus, it is overall rejected.

With respect to the effect size—which is important because the p-value can be distorted by large samples—only hypothesis H3 is considered to have practical relevance. The effect size of all other hypotheses indicates that the difference between the groups, although significant, has no relevance.

Discussion

Hypothesis 1: Book reviewers do not feel to obliged to write a positive review in return for the free copy. Even more, positive publicity might be more relevant for other product types than books; in the latter case, even negative publicity might be valuable.

Hypotheses 2/3: Book reviewers might not be aware of the possible danger of adding a disclosure statement; thus, the underlying assumption of „self-fulfilling prophecy" can be rejected. Still, an explanation for the confirmation of H3 could be the aforementioned *norm of reciprocity*.

Hypothesis 4: Reviewer motivations, while certainly different for incentivized and non-incentivized reviewers, are not reflected by extreme star ratings.

Hypothesis 5: It can be assumed that book reviewers develop an uniform writing style in reviews. Therefore, it does not make a difference whether the review is published shortly after the product experience or not.

Conclusion

Do incentivized book reviews show signs of influence if the reviewer received a free book copy? ?

Incentivization indeed impacts the contents of book reviews, but the only form of impact that has been found is an influence on review elaborateness (in terms of review length). At the same time, book reviews do not differ with respect to positivity, complexity, extremeness, and objectivity. !

However, the phenomenon of „influence" needs further investigation because there might be more dimensions than just the five considered in this thesis. Also, a conclusion such as „longer reviews are influenced" is abridged.

Limitations
The findings are only valid with respect to this study's product type, genre, language, reviewing platform, and temporal limitation.

Further Research Perspectives
- different formalisation and operationalisation of the concepts
- repeat analysis with sentence-based data
- try to avoid misclassification of NIRs
- use different NIR sample
- avoid biases by analyzing intra-reviewer or intra-book differences
- analyze a different genre
- derive hypotheses from book market-specifics

etc.

References
[1] Börsenverein d. Deutschen Buchhandels, Abt. Marktforsch. u. Statistik (Ed.). (2024). Buch und Buchhandel in Zahlen 2024: Zahlen, Fakten und Analysen zur wirtschaftlichen Entwicklung. MVB, here p. 85.
[2] Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? Journal of Interactive Marketing, 18(1), 38–52. DOI.
[3] Harper Collins (ed.) (2024). Incentivization. Collins Dictionary. URL.
[4] die medienanstalten – ALM GbR (ed.) (2018). Leitfaden der Medienanstalten – Werbekennzeichnung bei Online-Medien 2018. URL.
[5] Rebora, S., Messerli, T., & Herrmann, J. B. (2022). Towards a Computational Study of German Book Reviews—A Comparison between Emotion Dictionaries and Transfer Learning in Sentiment Analysis. DHd2022: Kulturen des digitalen Gedächtnisses. DOI.
[6] Kim, S. J., Maslowska, E., & Tamaddoni, A. (2019). The paradox of (dis)trust in sponsorship disclosure: The characteristics and effects of sponsored online consumer reviews. Decision Support Systems, 116, 114–124. DOI, here p. 116.
[7] Gouldner, A. W. (1960). The Norm of Reciprocity: A Preliminary Statement. American Sociological Review, 25(2), 161. DOI, here p. 170.
[8] Jacobs, A. M. (2019). Sentiment Analysis for Words and Fiction Characters From the Perspective of Computational (Neuro-)Poetics. Frontiers in Robotics and AI, 6, 53. DOI.

JOHANNES GUTENBERG UNIVERSITÄT MAINZ
HOCHSCHULE MAINZ UNIVERSITY OF APPLIED SCIENCES