

# Admiration and Frustration: A Multidimensional Analysis of Fanfiction

Mia Jacobsen<sup>1,\*</sup>, Ross Deans Kristensen-McLachlan<sup>1,2</sup>

<sup>1</sup>Center for Humanities Computing, Aarhus University, Denmark

<sup>2</sup>Department of Linguistics, Cognitive Science, and Semiotics, Aarhus University, Denmark

## Abstract

Why do people write fanfiction? How, if at all, does fanfiction differ from the source material on which it is based? In this paper, we use quantitative text analysis to address these questions by investigating linguistic differences and similarities between fan-produced texts and their original sources. We analyze fanfiction based on *Lord of the Rings*, *Harry Potter*, and *Percy Jackson and the Olympians*. Working with a corpus of around 250,000 texts containing both fanfiction and sources, we draw on Biber's Multidimensional Analysis [4], scoring each text along six dimensions of functional variation. Our results identify both global and community-based preferences in the form and function of fanfiction. Crucially, fan-produced texts are found not to diverge from their source material in statistically meaningful ways, suggesting that fans mimic the writing style of the original author. Nevertheless, fans as a whole prefer stories with less focus on narrative and greater emphasis on character interactions than the source text. Our analysis supports the notion proposed by qualitative studies that fanfiction is motivated both by admiration for and frustration with the canon.

## Keywords

quantitative text analysis, fanfiction, multidimensional analysis, style and genre, statistical modelling

## 1. Introduction

In 1992, Henry Jenkins published a seminal book in fan research, *Textual Poachers*. Contrary to the received opinion that fan cultures comprise misfits, degenerates, and mindless consumers to be ridiculed, fans are “active producers and manipulators of meaning” [16]. Drawing on the concept of textual ‘poaching’ developed by Michel de Certeau [7], Jenkins argues that fans actively transform the consumption of a given media into a participatory culture. This behavior is a product of both adoration and frustration with the media, motivating fans to explore and articulate the ways in which the narrative was unsatisfying and could be ‘salvaged’. A central aspect of this metaphorical salvation is the creation and dissemination of cultural products within fan communities - also known as *fandoms*. These products cover a wide range of media, such as videos, art, and playlists. For many people outside of these communities, though, the prototypical example of these fan productions is likely to be written texts. These examples of fan writing are both novel and derivative, allowing fans to create narratives of their favorite

---

CHR 2024: Computational Humanities Research Conference, December 4–6, 2024, Aarhus, Denmark

\*Corresponding author.

✉ miaj@cas.au.dk (M. Jacobsen); rdkm@cc.au.dk (R. D. Kristensen-McLachlan)

🔗 <https://orcid.org/0009-0003-3720-3418> (M. Jacobsen); <https://orcid.org/0000-0001-8714-1911>

(R. D. Kristensen-McLachlan)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

media in opposition to the original creators' intentions. These writings, also known as *fanfiction* (or *fanfic*), were popularized within fan circles mainly through the creation of fan zines in the 1950's and 1960's [15], but are now much more commonly posted to online fanfiction sites, with notable examples being Fanfiction.net and Archive of Our Own (AO3) [10].

Fanfiction as a textual genre is commonly defined as stories involving characters and worlds taken from a preexisting storyworld [24, 27, 2]. However, this definition neglects to mention the influence the fan communities have on fanfiction as a medium, despite individual fictions being rooted in specific fandoms. Indeed, fanfiction writers themselves report that the community is the main motivator for producing and disseminating these texts [1]. Moreover, the norms within the communities mean that writers often receive supportive feedback on their fanfiction stories [8]. This feedback is often responded to by later incorporation of the wishes of commenters into the fanfiction [9, 5]. In this way, fanfiction and the fan community have a reciprocal, even dialectical, relationship. The fans feed their wishes into the fanfiction, and the fanfiction becomes the main medium for co-creation and distribution of the norms and values within the community. As Busse also puts it in *Framing Fanfiction*: "fan fiction loses its meaning if removed from its context. Fan fiction thus offers insight into the fan community – its conversations, its tropes, and its members' discussions and concerns." [6]

Any proper consideration of fanfiction must therefore take into consideration both the linguistic structure of the texts *and* the unique context provided by the process of co-production within fan communities. In this paper, we hence set out to address the following research question: what does the linguistic structure of fanfiction tell us about the motivations and concerns of the fandoms that produce them? Is fanfiction as a genre monolithic or do different communities have different preferences?

### 1.1. Related Works

It is clear that the structure of individual fandoms and the texts they produce are potentially of great interest, insofar as they provide insight into the genesis of *interpretive communities* [12]. Nevertheless, there is a relative scarcity of scholarly research into fandoms and fanfiction, despite the sheer volume of data available online. Fanfiction research has traditionally been developed from a qualitative and ethnographic perspective [2]. However, given the volume of online text available through platforms such as AO3 and the prominence of these texts in online spaces, there is an increasing interest in the computational analysis of fanfiction [29].

Computational studies of fanfiction can be split into two main groups: those interested in the traits of popular fanfiction; and those interested in the character and gender dynamics of fanfiction more generally. For example, previous studies have found both gender and character disparities in fanfiction texts, with fanfiction being more likely to deprioritize the main characters in favor of the secondary characters and devote more attention to female characters [18]. Another study have found fanfiction pertaining to Greek mythology to be more likely to contain violence when the story is about a heterosexual couple compared to other couple constellations [19].

Concerning the textual features of successful or popular stories, these fanfics are found to have a simpler syntactic structure, a plainer writing style, but also a wider vocabulary [17, 20]. The features that pertain to direct speech are also more prevalent in popular fanfics compared

to other fanfics [17]. A different study comparing the emotional arcs and characters graphs of fanfiction found that fans preferred fanfictions with emotional arcs that were dissimilar to the source text's emotional arc, indicating a preference for stories with a different turn of events [26]. On the other hand, from the perspective of character networks, no clear global preference can be found regarding similarity or dissimilarity to the networks found in the source texts [26].

## 1.2. Multidimensional Analysis

Something which is currently missing from these quantitative analyses is an explicit linking between form and function. In other words, the distribution of individual linguistic features does not strictly tell us anything meaningful about the structure of the texts *as* texts or about the effect of these quantitative differences on readers. While this issue of interpretation is arguably a more fundamental problem in quantitative text analysis, it is nonetheless true that specific methods of analysis lend themselves more naturally to less speculative kinds of interpretation. This is particularly true of fields such as corpus stylistics, where a range of inferential statistics and null-hypothesis tests are integrated with stylistic analysis of authorial choices to explain variation in texts and across corpora.

One such approach is Biber's Multidimensional Analysis (MDA) [4]. MDA has been widely adopted across multiple different textual registers and genres and across multiple different languages. The core component of MDA involves analyzing the distribution of specific grammatical-semantic linguistic features which are argued to be *functionally motivated*. These features are grouped into metacategories, allowing us to describe the structure of texts along a number of *dimensions of variation*. This gives us a way of comparing the linguistic structure of texts and to explain what that variation means in terms of what those features *do* in a text.

To our knowledge, there are no studies that use MDA in the study of fanfiction. This study thus takes a novel approach to the study of fanfiction, one focused on the usage of linguistic features across text types to investigate the motivation and desires of fanfiction readers and writers

## 2. The Corpus

We chose to work with three particular fandoms, each of which are based on literary works of fantasy. Specifically, J.K. Rowling's *Harry Potter* series (HP), Rick Riordan's *Percy Jackson and the Olympians* series (PJ), and J.R.R. Tolkien's *Lord of the Rings* trilogy (LOTR) were chosen. These three groups constitute some of the biggest fandoms based on literary, fantasy novels. Limiting the scope in this way arguably limits the generalizability of our study but it also allows for a clearer comparison between fanfiction and source text, as well as a more controlled comparison across fandoms. We prioritized the robustness of the comparisons specifically because the pre-existing literature in the field is so limited.

The corpus of fanfiction was collected from the online fanfiction site AO3. This particular site is one of the largest repositories of fanfiction with over 13 million works and simultaneously functions as an online archive for fanfiction sites which no longer exist such as LiveJournal [28].

**Table 1**  
Corpus summary

	Harry Potter	Percy Jackson	LOTR	Total
Fanfiction	224,039	17,823	12,523	254,385
Source texts	202	69	84	355

On AO3, fanfics are split into fandoms through the use of tags. Authors add fandom tags to their fanfics when posting them, and they are used to specify which fictional universes a given story is connected to. A team of volunteers called *tag wranglers* make sure that tags are appropriately aggregated, so that, for example, fanfics with misspelled tags are additionally tagged with the correct one [23]. These tags were therefore used for retrieving the fanfiction texts, specifically, the tags “Harry Potter – J. K. Rowling”, “Percy Jackson and the Olympians – Rick Riordan”, and “Lord of the Rings – J. R. R. Tolkien”, since these tags pertain to the original, literary installments of the source texts. Although the stories might also be tagged with other fandoms such as “Harry Potter (movies)”, they are at least to some degree about the original texts, denoted by the use of these tags. Additionally, only fanfics written in English and fanfics which had no crossovers – meaning characters or worlds from other fandoms – were included. Besides that, all maturity ratings, lengths, and completion statuses were included in the scraping process.

We modified an already existing web scraper to collect data from AO3,<sup>1</sup> with minor adjustments adapted to fit the current study. The texts themselves were scraped over the course of late 2023 through early 2024. The latest text included in this study was published January 3rd, 2024, and the earliest is stated as being published on January 1st, 1950 (although this is almost surely due to errors in the archival process). Along with the texts, we also collected engagement metadata from AO3, such as the number of “hits” and “kudos” - i.e. how many times the story had been read and liked by the community. The fanfics were scraped in accordance with AO3’s terms of service and stored in compliance with GDPR.

In line with the tags that were used to collect the fanfiction, we limited our data to only the “core” texts in the original series. For LOTR this meant that only the three books were included: *The Fellowship of the Ring*, *The Two Towers*, and *The Return of the King*. *The Silmarillion* and *The Hobbit* were excluded even though they are also written by Tolkien and take place in the same fictional universe, since the use of fandom tags had effectively excluded fanfics pertaining to only *The Silmarillion* or only *The Hobbit*. For PJ, the study only includes the five original books: *The Lightning Thief*, *the Sea of Monsters*, *the Titan’s Curse*, *the Battle of the Labyrinth*, and *the Last Olympian*. Finally, HP included only the original seven books: *The Philosopher’s Stone*, *the Chamber of Secrets*, *the Prisoner of Azkaban*, *the Goblet of Fire*, *the Order of the Phoenix*, *the Half-Blood Prince*, and *the Deathly Hallows*.

Before feature extraction and modeling, some data cleaning measures were implemented, these are detailed in Appendix A.1. A summary of the number of texts can be seen in Table 1.

<sup>1</sup><https://github.com/radiolarian/AO3Scraper>

**Table 2**

Summary of dimensions of variation established using MDA

	Summary	Short Description
Dimension 1	Involved versus Informational	<i>Informational</i> : Dense and careful information integration. <i>Involved</i> : Verbal style with a focus on the here-and-now.
Dimension 2	Narrative Concern	Distinguishes between texts with a narrative focus from others
Dimension 3	Context-(in)dependent referents	<i>Context-dependent</i> : Receiver must use context to infer what time and place is being referred to. <i>Context-independent</i> : The referents in the text are made explicit and thus not dependent on the context
Dimension 4	Overt expression of persuasion	The degree to which the sender's opinion is overtly expressed and/or overt attempts to persuade the receiver are made
Dimension 5	Abstract style	Distinguishes between informational discourse that is abstract and technical from informational discourse that is not
Dimension 6	On-line information elaboration	Indicates whether the information presentation is made carefully or is more fragmented

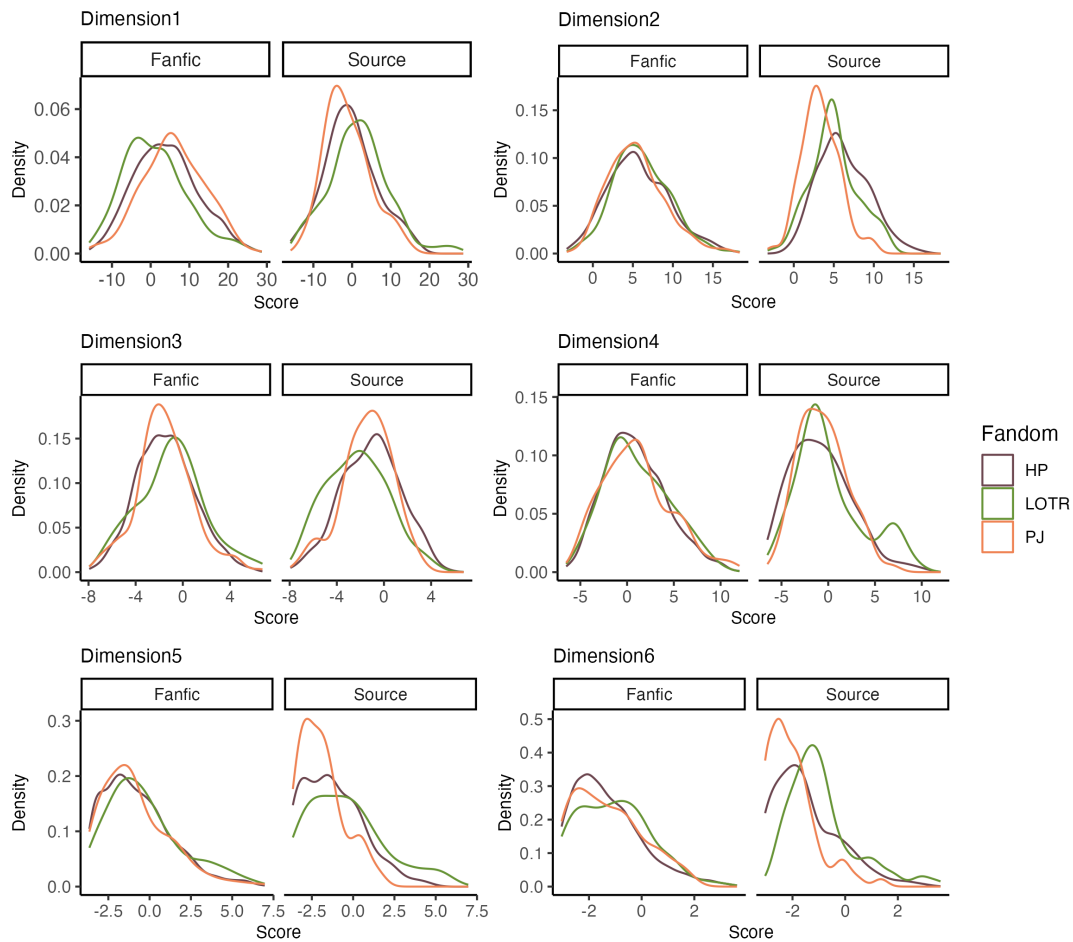
### 3. Experiment

#### 3.1. Feature Extraction

We extracted MDA features using the Multidimensional Analysis Tagger (MAT) [21]. The tagger creates grammatically annotated version of the texts by using a combination of the Stanford Tagger<sup>2</sup>, as well as a series of rules for identifying the patterns of linguistic features described in Biber's original study [4]. This allows the user to input either a single text or a whole corpus and receive both a tagged version of the text(s) and the different dimension scores for that text. In other words, the MAT scores each of the texts in the new register on the already established dimensions of variation within the English language. This means that the corpus of texts provided by a user is described relative to other prominent registers in English.

Each text is thus given a score for each of the six dimensions of variation. The different dimensions and their interpretations can be seen in Table 2. For a full description see [4, 21].

<sup>2</sup><https://nlp.stanford.edu/software/tagger.html>



**Figure 1:** Distribution of dimension scores across text types and fandoms.

### 3.2. Statistical Models

Based on the output from the MDA performed by the MAT, we conducted two statistical analyses. For both analyses we used linear mixed effects models. This type of statistical model has the advantage of accounting for two of the study’s most prominent challenges: Imbalance of sample sizes and repeated measures. Both of these challenges are addressed in the model formulation. Through the specification of fixed and random effects, the hierarchical structure of the data is built into the model. It thus accounts both for repeating authors and produces robust results when faced with imbalanced datasets [11, 14].

The first model was set up to test whether there is a statistical difference between fanfiction and source texts when it comes to the dimension scores extracted by the MAT. For each dimension, a linear mixed effects model was created which sought to predict the dimension scores from the text type (fanfiction / source text) and the fan group (HP/LOTR/PJ). A random intercept for author was included to control for the repetition of authors in the dataset. Because of

the great imbalance in the number of source texts compared to fanfic texts, we applied a set of weights to down-weight the fanfics and up-weight the source texts. These are specified in Appendix A.4. The model for this analysis is described as follows:

$$Dimension \sim text\ type + fandom + (1|author) \quad (1)$$

The second analysis included only the fanfiction texts and investigated differences in how readers respond to the different dimensions across fandoms. First, we defined an engagement metric inspired by Pianzola et al [23]. The engagement metric is computed as the number of kudos (i.e., likes) divided by the number of hits times 100 to get a percentage. In other words, it can be thought of as the percentage of people who read a fanfic and also decided to give it a like. Despite this metric not accounting for re-reads and updates,<sup>3</sup> we deemed it to be suitably representative as an engagement metric. We again created a linear mixed effects model for each dimension. These models sought to predict the dimension score based on an interaction between the engagement metric and the fan group, with the standardized word count and publishing date included as control variables. Similarly to the first analysis, a random intercept for author was included. Due to the large amount of data in each group, it was not deemed necessary to include weights in this model. The formula for the second analysis was as follows:

$$Dimension \sim engagement * fandom + published\ date + word\ count + (1|author) \quad (2)$$

After fitting the models, the assumptions of linear mixed effects regression were checked, and they were deemed to not be violated (see Appendices A.2 and A.3).

## 4. Results

The overall distribution of scores across each dimension can be seen in Figure 1. Most strikingly, these distributions are closely aligned across all dimensions for all fanfics, indicating a marked uniformity of linguistic style across each of the dimensions of variation. However, there are subtle differences which can be teased apart through the mixed effects models described above. The results from the two models are presented in Tables 3 and 4 respectively.

Table 3 shows that there is no significant difference between fanfics and source texts across the different dimensions of variation. Instead, there are only differences in these scores between individual fandoms. With regards to *informational/involved discourse* (D1), LOTR has the greatest degree of informational discourse and PJ has the greatest degree of involved discourse, while HP is in between the two groups. Additionally, LOTR has the greatest *narrative concern* (D2) of the three groups as well as the greatest *context-independence* (D3), while PJ has the least *narrative concern* (D2) and the greatest *context-dependence* (D3). Together, these three dimension indicate that fanfiction authors might be mimicking the style of the original author. The prevalence of *abstract style* (D5) also support this interpretation, since LOTR has the greatest degree of abstract style, while PJ has the least abstract style among these three groups. These differences in fandoms for D1, D3, and D5 are similarly found in the second analysis (see Table 4).

<sup>3</sup>Users can open a fanfic multiple times - when there are updates, for instance - but they can only like it once.

**Table 3**  
Estimates for model (1) for each dimension of variation

Dimension 1		$\beta$	SE	t-value	p-value
	Text type	-3.12	2.44	-1.28	0.20
	LOTR	-2.30	0.08	-28.78	<0.001*
	PJ	1.47	0.064	23.16	<0.001*
Dimension 2		$\beta$	SE	t-value	p-value
	Text type	-0.58	1.2	-0.5	0.62
	LOTR	0.24	0.037	6.29	<0.001*
	PJ	-0.45	0.03	-15.25	<0.001*
Dimension 3		$\beta$	SE	t-value	p-value
	Text type	-0.32	0.71	-0.44	0.66
	LOTR	0.14	0.025	5.71	<0.001*
	PJ	-0.23	0.02	-11.3	<0.001*
Dimension 4		$\beta$	SE	t-value	p-value
	Text type	-1.36	0.98	-1.38	0.17
	LOTR	-0.033	0.034	-0.96	0.34
	PJ	0.25	0.028	9.22	<0.001*
Dimension 5		$\beta$	SE	t-value	p-value
	Text type	-0.46	0.57	-0.81	0.42
	LOTR	0.13	0.02	6.45	<0.001*
	PJ	-0.095	0.016	-5.76	<0.001*
Dimension 6		$\beta$	SE	t-value	p-value
	Text type	-0.19	0.36	-0.53	0.6
	LOTR	0.098	0.013	7.66	<0.001*
	PJ	-0.081	0.01	-7.89	<0.001*

In contrast to the other dimensions, the model for *overt expression of persuasion* (D4) finds no difference in scores between HP and LOTR, but PJ texts generally have a greater degree of overt persuasion. There is no immediately apparent reason for this pattern. One interpretation concerns the domain of authorial point-of-view and modality [25]. While it is not possible to explore this result in detail in this paper, it does suggest that future research may need to pay greater attention to the rhetorical aspects of fanfiction than has previously been afforded to the genre.

When it comes to *on-line information elaboration* (D6) another unexpected pattern of findings emerges. We find that PJ is generally more careful and planned in its information presentation, whereas LOTR seems to have a more fragmented information presentation. Since both *abstract style* (D5) and *informational/involved discourse* (D1) indicate the complete opposite pattern, it is unusual that PJ is significantly more planned than LOTR. The distributions of dimension scores illustrated on Figure 1 also show some unusual patterns when it comes to *on-line information elaboration* (D6). When looking at Figure 3, which shows the result of the second analysis, the



**Table 4**  
Estimates for model (2) for each dimension of variation

Dimension 1	$\beta$	SE	t-value	p-value
LOTR	-2.59	0.16	-15.90	<0.001*
PJ	1.76	0.14	12.18	<0.001*
engagement	0.13	0.0057	22.06	<0.001*
engagement:LOTR	0.027	0.021	1.28	0.2
engagement:PJ	-0.047	0.020	-2.42	<0.05*
Dimension 2	$\beta$	SE	t-value	p-value
LOTR	0.15	0.076	1.94	0.052
PJ	-0.66	0.068	-9.77	<0.001*
engagement	-0.048	0.0027	-18.07	<0.001*
engagement:LOTR	0.032	0.0098	3.24	<0.01*
engagement:PJ	0.052	0.0092	5.67	<0.001*
Dimension 3	$\beta$	SE	t-value	p-value
LOTR	-1.43	0.051	2.81	<0.01*
PJ	-0.26	0.046	-5.64	<0.001*
engagement	-0.0037	0.0018	-2.05	<0.05*
engagement:LOTR	-0.0012	0.0067	-0.18	0.86
engagement:PJ	0.0016	0.0063	-0.25	0.8
Dimension 4	$\beta$	SE	t-value	p-value
LOTR	0.22	0.067	0.035	0.97
PJ	0.046	0.061	0.77	0.44
engagement	0.040	0.0024	16.71	<0.001*
engagement:LOTR	-0.0056	0.0088	-0.64	0.52
engagement:PJ	0.038	0.0083	4.59	<0.001*
Dimension 5	$\beta$	SE	t-value	p-value
LOTR	0.0024	0.041	5.29	<0.001*
PJ	-0.10	0.038	-2.74	<0.01*
engagement	-0.011	0.0015	-7.63	<0.001*
engagement:LOTR	-0.0098	0.0054	-1.80	0.071
engagement:PJ	-0.013	0.0051	-0.26	0.80
Dimension 6	$\beta$	SE	t-value	p-value
LOTR	0.097	0.026	3.81	<0.001*
PJ	-0.074	0.023	-3.21	<0.01*
engagement	-0.0082	0.00092	-8.91	<0.001*
engagement:LOTR	-0.0028	0.0033	-0.85	0.39
engagement:PJ	-0.037	0.0031	-1.17	0.24

findings are again worth questioning. The y-axis for D6 is on a tiny range (from -1.1 to -1.4), and the confidence intervals of the regression lines are quite wide. As D6 was dropped in later iterations of the MDA [3], we would argue that omitting it within the context of this study

makes more sense than including it.

Table 4 describes the relationship between engagement and dimension scores across fandoms and shows a yet more nuanced picture. We found significant interaction effects for *informational/involved discourse* (D1), *narrative concern* (D2), and *overt expression of persuasion* (D4). This means that for these three dimensions, there is a significant difference in how fans respond to the dimension scores across groups. For *context-(in)dependent referents* (D3), *abstract style* (D5), and *on-line information elaboration* (D6) there was no differences in how fans across groups respond to these linguistic features.

For *informational/involved discourse* (D1), there is a positive relationship between engagement and dimension scores, meaning that across fandoms there is a preference for fanfics that are more involved. Although this effect persists in all fandoms, it is smaller for the PJ fandom, while there is no significant difference in how fans from HP and LOTR respond to this dimension.

For *narrative concern* (D2), the difference between HP and LOTR is no longer present when purely looking at the main effect of fandom, but PJ has a significantly lower degree of narrative concern compared to the other groups. For the engagement score there is a negative main effect, meaning that less narrative concern is associated with greater engagement. This effect is different across groups, with HP having a stronger effect than the other two groups.

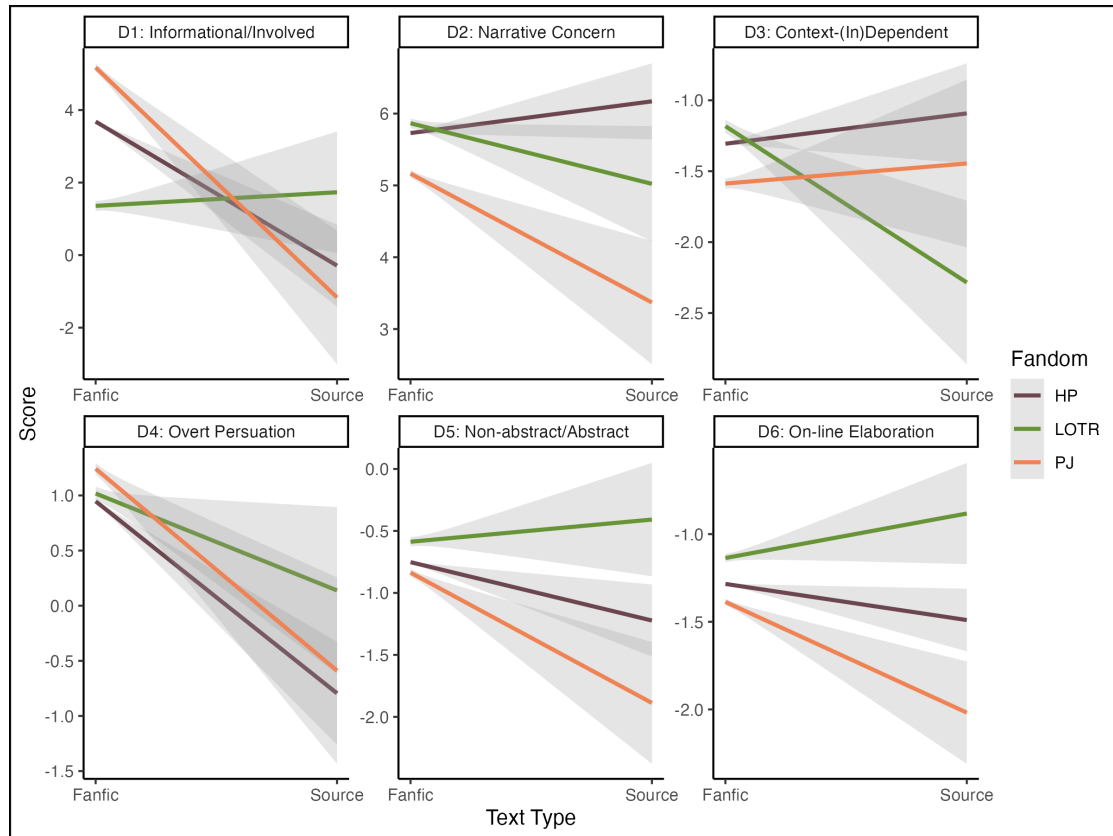
The differences in *expression of persuasion* (D4) across fandoms also disappears in the second analysis. Nevertheless, there is a significant, positive main effect of engagement on overt persuasion across groups, and a significant, positive interaction effect of engagement for PJ fanfics specifically. In other words, across fandoms there is a general preference for more overt expression of persuasion, but this effect is especially strong for PJ.

For both *context-(in)dependent referents* (D3) and *abstract style* (D5) there is an association between greater engagement and a lower dimension score, meaning that across the three fandom groups, they all respond with a preference for texts that have a greater context-dependence and less abstract information. It is worth noting that the effect for D3 is vanishingly small ( $\beta = 0.004$ ), but that taken together with the other findings, it could again illustrate that fans prefer texts that are more here-and-now oriented and less technical, thus exhibiting the same patterns as seen throughout this analysis.

## 5. Discussion

What do these results mean for the study of fanfiction and the fandoms that produce them? We consistently find no statistical difference in dimension scores between fanfiction and their original source texts, only between fan groups. Some might argue that this is unsurprising since both fanfiction and their source texts fall into the more general category of 'fiction', which limits the actual variation that might occur compared to other registers. This view neglects, however, that different genres of fiction have been found to differ in MDA [4, 3]. Additionally, despite finding no variation between the two text types, the analysis does pick up on arguably more subtle differences in writing style between the original authors as exhibited by the differences across fandoms.

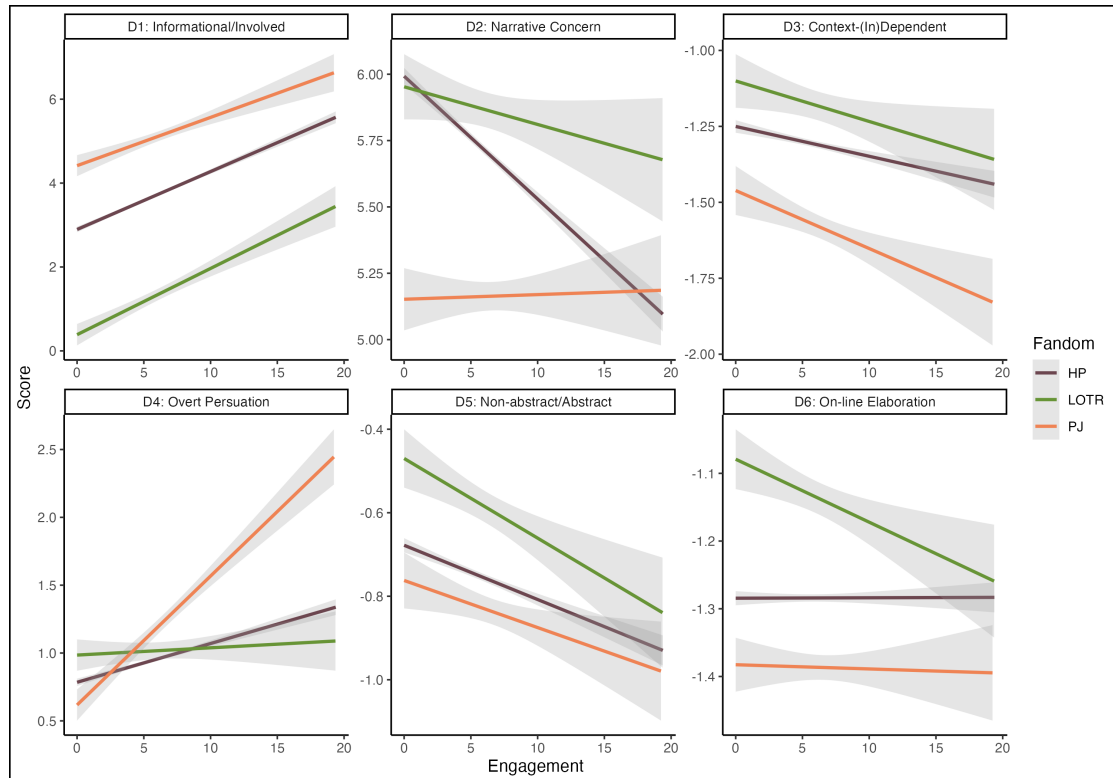
Despite the lack of a difference between the text types, fanfiction does seem to be more ho-



**Figure 2:** Mean dimension scores across text types and fandoms.

mogeneous than expected. Although previous research has found fans to be a heterogeneous group, fanfiction exhibits a quite consistent linguistic style as compared to the source texts, as seen in the distributions of dimension scores in Figure 1 and the regression analysis illustrated in Figure 2. Despite the plot indicating a difference in means between the fanfics and source texts, the source texts' confidence intervals are much wider than the fanfics', which probably drives the statistically insignificant results. The analysis thus indicates that fanfiction in general has a distinct style that is more consistent than the styles across source texts, but that this style within communities is influenced by the writing style of the source text. Since we did not find a statistical difference between fanfiction and source texts, further research is needed to identify with certainty if this fanfiction style exists and how it integrates the style of its source material.

Our second experiment foregrounds the way different communities of fans respond to the prevalence of different textual features in their fanfics. By investigating the degree to which fans appreciate the prevalence of different linguistic features in the texts, we find that fans write fanfiction that generally looks similar across dimensions and mimic the style of the original author, but they might not appreciate the same traits. All the fandoms studied prefer less narrative concern, less abstract information, more conversational style, and discourse focused



**Figure 3:** Regression lines between engagement metric and dimension scores across fandoms.

on the here-and-now. These preferences illustrate that although fanfiction imitates the original writing style of the source author, fans across groups still have a preference for stories that are less imitative and more focused on the core aspects of fanfiction, namely character interaction and emotional experiences [2, 24, 15].

When looking at the local, community-specific preferences as illustrated on Figure 3, LOTR fans seem to prefer greater *narrative concern* (D2) as compared to the other two fandoms. This preference is, simultaneously, a prominent aspect of the style of writing that sets LOTR apart from the others. In contrast, PJ fans have no preference for narrative concern, but strongly prefer greater *overt expression of persuasion* (D4) when compared to the other two fandoms. Again, PJ texts in general were found to score higher on this dimension - something that put those texts apart from the others. HP, the fandom in the middle of the spectrum, shows a strong preference for less narrative concern. One interpretation could be that these three fandom groups have distinct preferences for the degree of narrative in their fanfics. HP fans thus write fanfics with a greater variety in narrative concern but have a strong tendency to prefer works with less narrative. Meanwhile, LOTR fans are more inclined to prefer greater narrative concern and thus also write works that fit this – a trait likely inherited from their source text. Lastly, for PJ fans, the flat regression line visible on Figure 3 could be an indication that they do not respond to narrative concern at all. In other words, they appreciate fanfics with a tendency

for either end of this dimension, but writers are more inclined to write with less narrative concern, which could also be an inherited trait from their source text.

So, while global preferences do exist for involved discourse, non-abstract information, and here-and-now focus, community-specific preferences seem to arise from the linguistic features that set specific fandoms apart from other texts. This analysis supports the idea that fans write fanfiction due to both admiration and frustration with the source material. The admiration is seen in the imitation of the original author's style, whereas the frustration is seen in the preference for fanfics that break with the mold. This idea of admiration and frustration is exactly the argument put forth by Jenkins [16] and later echoed through other studies [2, 24].

## 6. Limitations

One limitation concerning the data collection is the fact that all the analyzed fandoms are based on fantasy novels. This was necessary due to the scope of the study where the simplicity of the study design and robustness of findings were prioritized. This made the inclusion of other types of source material not feasible. It does, however, impact the generalizability of these findings, which might only be present in fanfiction based on fantasy novels.

Another limitation concerns the comparison between the source texts and the fanfiction texts. All the included fandoms had prior to the scraping of the texts already been adapted to TV and/or film. As it was near impossible to make sure that fanfiction based on only the actual texts were included, all fanfiction stories which had the specified tags were included. This limits the robustness of the comparison, as fans might not have read the source texts before writing their fanfic, basing their fanfiction entirely on the adaptation. The idea of only including fanfiction based on the actual source texts is, however, not as sensible as it might seem. Fans are known to write fanfiction based on shows or stories they have not consumed themselves [24]. Thus, knowing whether a piece of fanfiction is based on just its source text, an adaptation, or even another fanfiction is impossible.

Finally, we operationalize engagement as a kudos/hits ratio meaning that fanfics are effectively 'punished' for being revisited multiple times. Moreover, the metric is quite crude, especially if the goal is to deeply understand reader preferences when it comes to fanfiction. However, as very little previous research has taken a computational approach to fanfiction reader appreciation, this study is a step in the direction of a more nuanced understanding of fanfiction as a phenomenon.

## 7. Conclusion

Despite the dynamic, dialectical co-production of fanfiction by specific fandoms, the resulting texts are not significantly different from their source material, focusing instead on mimicking the voice of the original author. While fans do mimic the voice of the original author, across communities we find that fans prefer fanfiction stories that are more conversational and here-and-now oriented, meaning a preference for fanfiction stories that are different than their source material. Fans appear to be more interested in character interactions than in plot. This trend is only to some degree, however, as the evidence also suggests that fans prefer the

linguistic features that set their source text apart from other groups of fans. Our analyses thus support the conclusion that fans write fanfiction both due to admiration and frustration with the source material – similar to what previous, qualitative studies have found.

Our study hence has a two-fold contribution. Firstly, it shows that these qualitative findings are replicated when taking a quantitative approach, thereby providing additional support for the reliability of these arguments. Secondly, our experiments illustrates how one can answer the *why* of fanfiction writing by inferring it from an analysis of the *how*. Specifically, we find that the imitation of the original author’s writing style could be an expression of admiration, while the greater appreciation for fanfiction stories that are less imitative and perhaps more generically fanfiction could be an expression of the tension, frustration, and resistance to the source material.

## Acknowledgments

Part of the computation done for this project was performed on the UCloud interactive HPC system, which is managed by the eScience Center at the University of Southern Denmark.

## References

- [1] K. Bahoric and E. Swaggerty. “Fanfiction: Exploring in-and out-of-school literacy practices”. In: *Colorado Reading Journal* 26 (2015), pp. 25–31.
- [2] J. L. Barnes. “Fanfiction as imaginary play: What fan-written stories can tell us about the cognitive science of fiction”. In: *Poetics* 48 (2015), pp. 69–82. DOI: <https://doi.org/10.1016/j.poetic.2014.12.004>.
- [3] D. Biber. “A typology of English texts”. In: *Linguistics* 27.1 (1989), pp. 3–43.
- [4] D. Biber. *Variation across speech and writing*. Cambridge: Cambridge University Press, 1988.
- [5] R. W. Black. “Language, culture, and identity in online fanfiction”. In: *E-learning and Digital Media* 3.2 (2006), pp. 170–184. DOI: <https://doi.org/10.2304/elea.2006.3.2.170>.
- [6] K. Busse. *Framing fan fiction: Literary and social practices in fan fiction communities*. Iowa City: University of Iowa Press, 2017.
- [7] M. de Certeau. *The Practice of Everyday Life*. Berkeley: University of California Press, 1984.
- [8] J. S. Curwood, A. M. Magnifico, and J. C. Lammers. “Writing in the wild: Writers’ motivation in fan-based affinity spaces”. In: *Journal of Adolescent & Adult Literacy* 56.8 (2013), pp. 677–685. DOI: <https://doi.org/10.1002/JAAL.192>.
- [9] S. Evans, K. Davis, A. Evans, J. A. Campbell, D. P. Randall, K. Yin, and C. Aragon. “More than peer production: Fanfiction communities as sites of distributed mentoring”. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. Portland, Oregon, 2017, pp. 259–272.

- [10] J. Fathallah. “Digital fanfic in negotiation: LiveJournal, Archive of Our Own, and the affordances of read–write platforms”. In: *Convergence* 26.4 (2020), pp. 857–873. doi: <http://doi.org/10.1177/1354856518806674>.
- [11] A. Field, J. Miles, and Z. Field. *Discovering Statistics Using R*. Sussex: Sage, 2012.
- [12] S. Fish. *Is There A Text in This Class*. Cambridge: Harvard University Press, 1980.
- [13] L. Hansen, L. R. Olsen, and K. Enevoldsen. “TextDescriptives: A Python package for calculating a large variety of metrics from text”. In: *arXiv preprint arXiv:2301.02057* (2023).
- [14] S. Huber, E. Klein, K. Moeller, and K. Willmes. “Comparing a single case to a control group—applying linear mixed effects models to repeated measures data”. In: *Cortex* 71 (2015), pp. 148–159.
- [15] A. Jamison. *Fic: Why fanfiction is taking over the world*. Dallas, Texas: BenBella Books, Inc., 2013.
- [16] H. Jenkins. *Textual poachers: Television fans and participatory culture*. London & New York: Routledge, 1992.
- [17] A. Mattei, D. Brunato, and F. Dell’Orletta. “The Style of a Successful Story: a Computational Study on the Fanfiction Genre”. In: *Computational Linguistics CLiC-it 2020*. Bologna, 2020.
- [18] S. Milli and D. Bamman. “Beyond canonical texts: A computational analysis of fanfiction”. In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Austin, Texas, 2016, pp. 2048–2053.
- [19] J. Neugarten and R. Smeets. “MythFic Metadata: Exploring Gendered Violence in Fanfiction about Greek Mythology”. In: (2023).
- [20] D. Nguyen, S. Zigmond, S. Glassco, B. Tran, and P. J. Giabbanelli. “Big data meets storytelling: using machine learning to predict popular fanfiction”. In: *Social Network Analysis and Mining* 14.1 (2024), p. 58.
- [21] A. Nini. “The multi-dimensional analysis tagger”. In: *Multi-dimensional analysis: Research methods and current issues* (2019), pp. 67–94.
- [22] J. Ooms and D. Sites. “cld2: Google’s compact language detector 2”. In: *Retrieved February 7* (2018), p. 2019.
- [23] F. Pianzola, A. Acerbi, and S. Rebora. “Cultural accumulation and improvement in online fan fiction”. In: *CEUR Workshop Proceedings*. Vol. 2723. CEUR-WS. org. Amsterdam, 2020.
- [24] S. Pugh. *The democratic genre: Fan fiction in a literary context*. Brigend: Seren, 2005.
- [25] P. Simpson. *Language, Ideology and Point of View*. London & New York: Routledge, 1993.
- [26] Z. Sourati Hassan Zadeh, N. Sabri, H. Chamani, and B. Bahrak. “Quantitative analysis of fanfictions’ popularity”. In: *Social Network Analysis and Mining* 12.1 (2022), p. 42.
- [27] B. Thomas. “What is fanfiction and why are people saying such nice things about it??” In: *Storyworlds: A Journal of Narrative Studies* 3 (2011), pp. 1–24. doi: <https://doi.org/10.5250/storyworlds.3.2011.0001>.

- [28] C. Tosenberger. “Mature poets steal: children’s literature and the unpublishability of fanfiction”. In: *Children’s Literature Association Quarterly* 39.1 (2014), pp. 4–27. doi: 10.1353/chq.2014.0010.
- [29] K. Yin, C. Aragon, S. Evans, and K. Davis. “Where No One Has Gone Before: A Meta-Dataset of the World’s Largest Fanfiction Repository”. In: *Proceedings of the 2017 CHI conference on human factors in computing systems*. Denver, 2017, pp. 6106–6110.



## A. Appendix

### A.1. Data Cleaning

In the usual MDA analysis pipeline, the first 400 words of a text are extracted and tagged for linguistic features. However, many posts made to AO3 are not written stories but are instead picture collages, playlists, poems, audiostories, or other kinds of fan creations. Additionally, fanfiction stories often have so-called author notes at the beginning and end of a chapter, which are not part of the story itself. In an effort to minimize artefacts but keep representativeness, fanfics with less than 600 words were excluded, and the MDA was run on the middle 500 words of each fanfic.

After the snippets had been extracted, we utilized the `textdescriptives` package [13] to assess the quality of the snippets. Specifically the quality pipeline component which calculates the quality of the text based on both heuristic quality metrics and repetitious text metrics was utilized. We used the default quality check settings and filtered out texts that did not pass this quality check. The quality check was implemented to make sure that only fanfics which could be described as written stories were included, and not posts such as lists or picture collages.

For the source texts, we also deviated from the typical approach in MDA. Since there, in some sense, were only three source texts, it would amount to too little data if only 500 words were extracted from each. Therefore, to have a comparable corpus of source texts that were tagged, we extracted 500 words every 5000 words for each of the full source texts.

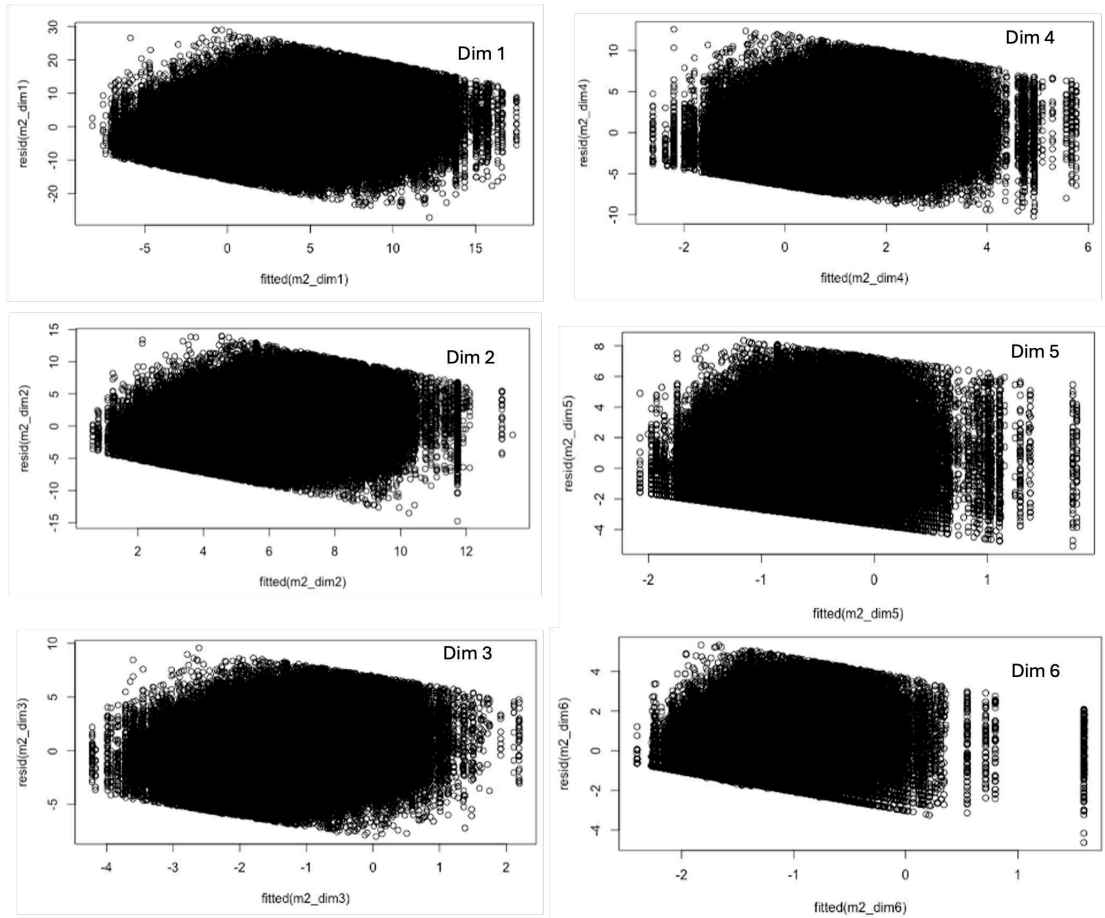
Before the statistical analysis, it was also deemed necessary to perform outlier removal. Not only is there a great variability in document length, but hits, kudos, and dimension scores also had quite extreme values. Outliers were defined as data points lying in the top and bottom 0.5% of a given distribution, in this case all of the dimension scores as well as hits, kudos, and word count. This method was chosen as it minimized the data that was excluded without compromising on the robustness of the statistical analysis compared to other methods (e.g., removing extreme outliers as defined by a boxplot). It also ensured that the cut-offs were as explicit as possible. A total of 7,026 fanfics and 67 source text snippets were excluded based on outlier removal. Additionally, as mentioned earlier, fanfics that did not pass the quality check from the `textdescriptives` package were excluded, which constituted 38,307 fanfics.

Finally, it was also necessary to perform additional language detection on the text snippets, as some texts were written in other languages than English. Using the `cld2` package in R [22], 319 fanfics were excluded as they were detected as a language other than English. Finally, since the second analysis included publishing date and word count as a control variables, the 12 fanfics that were set as published before January 1st, 2000 were removed, and the word count was standardized to aid in model convergence.

### A.2. Model (1) assumption checks

There are the following five assumptions of linear mixed effects modeling: Linear relationship between predictor and output variable(s), no multicollinearity, independence of data points, homoscedasticity, and multivariate normality.

The first three assumptions can be addressed jointly for all six models. Since the models compare the means of categorical groups, one might have run an ANOVA instead of linear



**Figure 4:** Test of homoscedasticity

regression. However, a mixed effects model was necessary due to repeated authors. Since ANOVA's can be conceptualized as a specific case of the general linear model, the linear relationship is built into the model formulation. Independence of data points is achieved through the random effects. Since all six models have the same predictors, we calculated the variance inflation factor (VIF) for the Dimension 1 model. Both predictors had a VIF of 1, indicating no multicollinearity.

For the fourth assumption, homoscedasticity, we have plotted the residuals against the fitted values for each model below. Although the plots show some downward trend in the residuals, due to the lack of a clear cone shape or other extreme heteroscedasticity the models were deemed to be not violating this assumption.

The fifth assumption is also tested by plotting the residuals in a qq-plot to ensure they are normally distributed. These are plotted below. All plots indicate that the residuals are generally normally distributed. As such, the assumption of multivariate normality was deemed not violated.

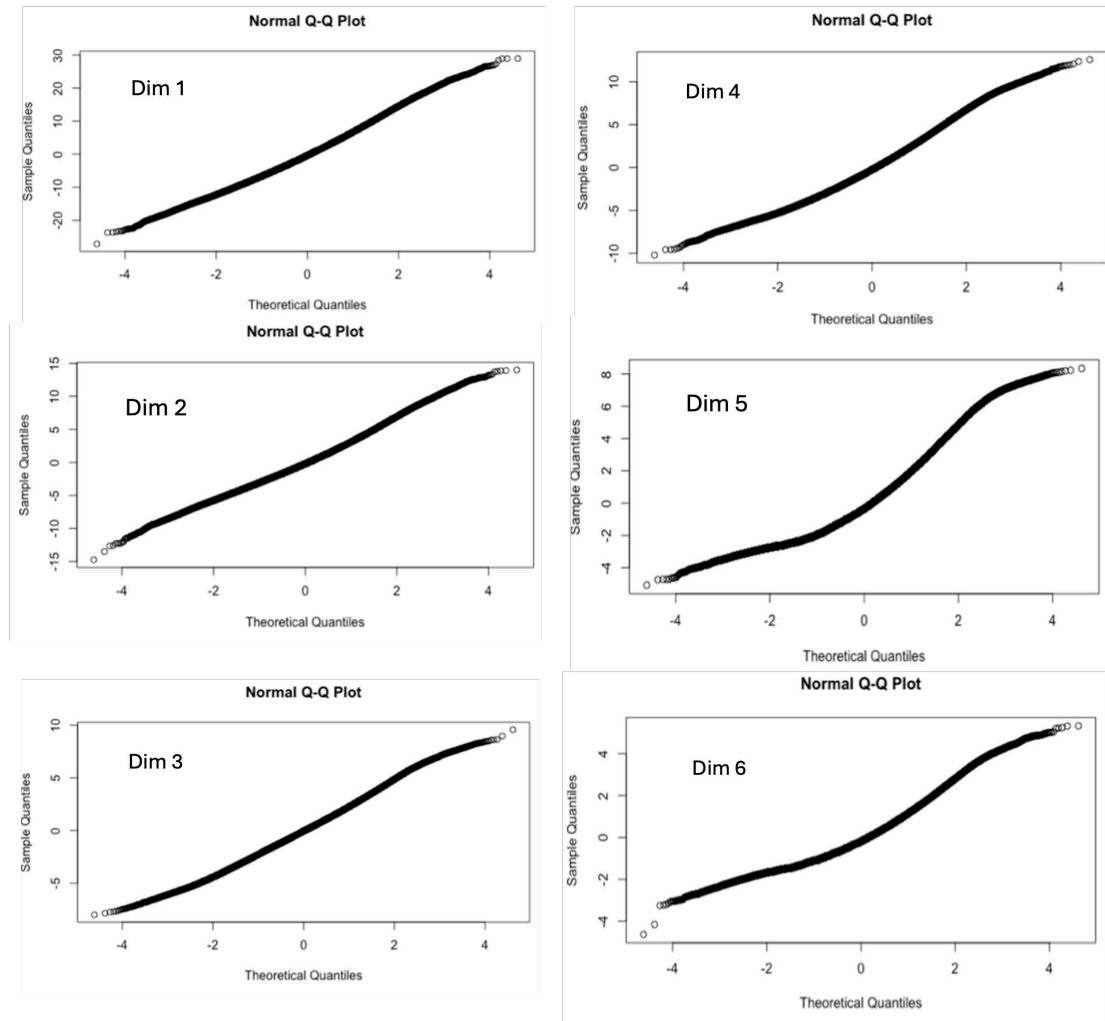
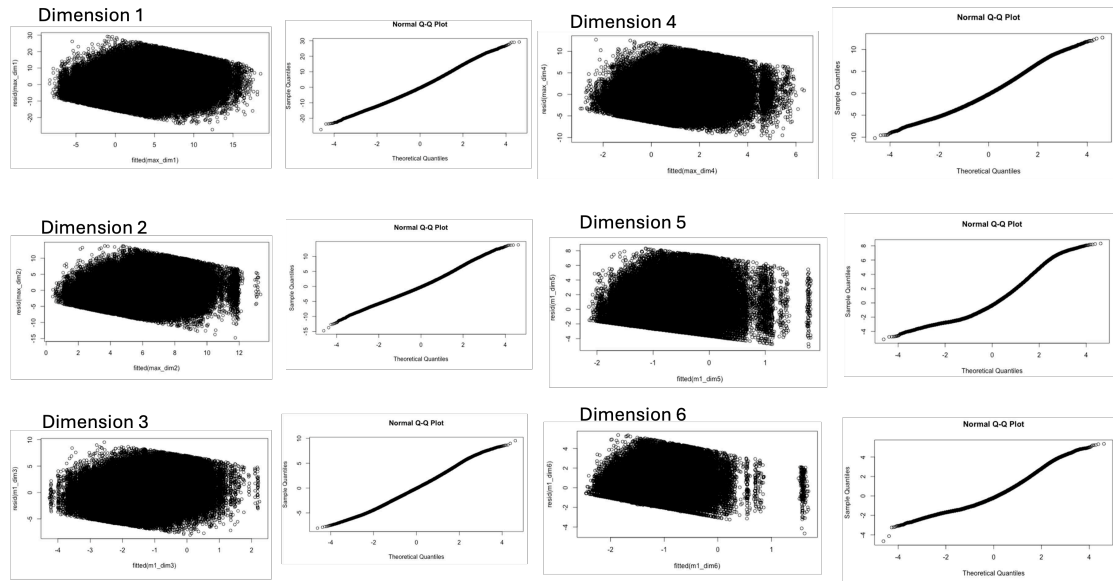


Figure 5: QQ-plot of the residuals for each model

### A.3. Model (2) assumption check

The first and third assumptions can be addressed at once for all models. Due to the sheer number of data points, it would be infeasible to plot them on a point plot to assess whether there is a linear relationship. Instead, we take special notice that all regression lines on figure 2 mostly have narrow confidence intervals, meaning that the line is quite confident in its placement. We would therefore argue that this assumption is not violated, however, it is worth exploring if other relationships than linear might explain the data better.

As with the previous model, the assumption of independence of data points is accounted for by the random effects. The VIF was calculated for the predictors in the model for Dimension 1. All VIF scores were below 5, meaning no violation of multicollinearity – except for the



**Figure 6:** Test of homoscedasticity and multivariate normality for each dimension score model

interaction terms. However, collinearity of interaction terms has been argued to be inevitable,<sup>4</sup> and as such, we deem the assumption not violated.

Finally, the tests for homoscedasticity and multivariate normality for each dimension's model are presented above. As with the previous models, no model residuals indicate any extreme deviations from homoscedasticity or normality, and as such the assumptions are not violated.

#### A.4. Weights applied to model (1)

Since there are 1000 times the number for HP fanfics as there are HP sources, and around 150 times the amount of PJ and LOTR fanfics as PJ and LOTR source texts, we accounted for this imbalance using the following weights. HP fanfics were weighted with  $1/\text{total number of fanfics}$  ( $= 0.00000393$ ). PJ fanfics and LOTR fanfics were weighted with  $1 \cdot 15 / \text{total number of fanfics}$  ( $= 0.0000597$ ), to account for the fact that there are around 15 times the number of HP fanfics as other fanfics. All source texts were weighted with  $1/\text{total number of source texts}$  ( $= 0.00282$ ). Although different weights can quite substantially change the outcome of the models, these weights were decided upon since they most accurately describe the different imbalances in the data.

<sup>4</sup>see: <https://www.statalist.org/forums/forum/general-stata-discussion/general/1359532-is-multicollinearity-between-interaction-terms-a-problem>