

Works in Progress • Digital Social Reading

Chapter 3. What digital social reading reveals about readers

Published on: May 03, 2021

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

3.1 Comments, reviews, and reader response

DSR practices can reveal many things about how people feel and think when they read. Comments on Wattpad or fanfiction websites, online book reviews, YouTube videos, Instagram posts, and Twitter discussions are a trove of evidence about the reading experience of a variety of readers. These sources enable reading research and literary studies to extend their scope of investigation beyond institutionalized reception and interpretation, as transmitted by book publishers, periodicals, and literary critics. For the first time in history, we have access to the verbalized response of millions of readers, an opportunity that can open new ways of understanding how lay readers choose what books to read and how they engage with them.

Research on reader response has started to show interest in online reading groups and book reviews to test theoretical hypotheses, to complement stylistic analysis, and as a source for ethnographic inquiry (Rehberg Sedo 2011b; Peplow et al. 2016; Nuttall 2017). A striking difference from first-generation reader response theory concerns the concept of reader: works of critics like Hans Robert Jauss, Wolfgang Iser, and Umberto Eco focused mainly on abstractions rather than on actual readers (Rose 1992; Whiteley and Canning 2017; Salgaro 2011; Willis 2021). Empirically-driven reader response studies, on the other hand, brought back the actual reader at the center of the analysis, by recognizing the relevance of their response for the study of literature (Miall 2018). Moreover, DSR data open new ways of doing research in computational literary studies and cultural analytics, disciplines which usually rely on publishing history when trying to reconstruct the reception of books (Bode 2018).

An additional advantage of studying digital reader response is that the ecological validity of the experimental method is not affected by interruptions of the reading experience or manipulations of the text due to methodological needs (Hall 2008). It is a “naturalistic study of reading” (Swann and Allington 2009). In historical perspective, glosses in ancient manuscripts and marginalia on printed books can offer a kind of insight similar to DSR comments and reviews, but their amount, cultural diversity, and circulation is very limited in comparison to the extent of the social interactions enabled by digital media (Jackson 2001; Kerby-Fulton, Hilmo, and Olson 2012). For instance, Wattpad publishes in more than 50 languages, has about 70 million readers, and nearly 300,000 writers from 35 countries take part every year in the largest writing competition in the world (Wattpad n.d.). Similarly, fanfiction platforms like AO3 and Fanfiction.net offer a huge quantity of data to study diversity in readership. Fanfiction

studies have grown into a thriving field of research but have been mostly applying qualitative approaches (Black 2008; Campbell et al. 2016; S. Evans et al. 2017; Alexander and Rhodes 2018), which offer precious insight but do not aim at a large-scale description of the variety, circulation, and reception of fanfiction. Among the few quantitative investigations of fanfiction platforms, interesting results have been achieved by Milli and Bamman (2016) – showing how fanfiction often deprioritizes main protagonists in comparison to canonical texts – by Yin et al. (2017) – who analyzed metadata to identify the most popular fanfiction genres – and by Aragon, Davis, and Fiesler (2019) – who showed how comments received from peers can help to improve writing skills.

Quantitative methods can be used to study reader response and understand how fiction is read nowadays, gaining unprecedented access to readers' interaction with texts and with other readers. Rowberry (2019) has shown how even the most sophisticated and pervasive sociotechnical reading ecosystems like Amazon are currently very limited for the study of reading practices. However, with the right methodology, big data can be fruitfully used to understand the new world of fiction reading (Pianzola, Rebora, and Lauer 2020; Rebora and Pianzola 2018), keeping in mind that working on aggregated data means focusing on collective readership. When analyzing aggregated readers' comments or reviews, we overlook individual differences in favor of collective trends, a choice that neglects the specificity of each individual reading, but can be helpful to outline informative patterns for literary history and education. The results of the analysis of reader response in large-scale environments can show the cognitive and emotional features of a collective act of reading, something similar to what Jonathan Rose (1992) imagined achieving by studying working-class autobiographies.

A general positive aspect to consider about large-scale DSR platforms is that the great quantity of comments and reviews alone can be effective in encouraging people to read a certain story or comment on it. Therefore, working on aggregated data can shed some light on the dynamics that motivate readers and elicit them to share their impressions with others. In the following sections I will present examples of how different digital tools can be used to study reader response.

3.2 Marginalia

Table 7. Wattpad features

DSR platform	Content macro-category	Relation to source text	Type of audience	Timeframe
Wattpad	Comment in the margin	Immediate	Global	During reading, self-paced, scheduled
	Comment in footer	Immediate	Global	During reading, self-paced, scheduled
	Rating	Immediate/external	Global	During/after reading, self-paced, scheduled
	List	External	Private, global	After reading, self-paced

Wattpad readers can write comments in the margin, immediately anchored to paragraphs, publicly sharing their thoughts and emotional reactions to specific parts of a story, not just to the whole book, as in the case of online reviews. This means that researchers can access detailed data about readers' response: comments that people write while they are reading, briefly (and spontaneously) interrupting their activity before continuing it. This "Immediate discussion within a community" (DSR type 1) is a kind of thinking-aloud or real-time data (M. Short et al. 2011; Canning 2017), which has never been available on a scale of millions of readers. Comments can also be written in the footer, anchored to a chapter. The timeframe can be both self-paced, when readers casually access stories, and scheduled, when they follow the updates of a story published in serial instalments and read it as soon as a new chapter is available. In the latter case, the interactive nature of serial publishing – also common on fanfiction platforms (Thomas 2011a) – activates readers' interest to participate in the story co-construction, by manifesting their emotional engagement with characters and expressing their desired plot development.

Other types of content generated by readers are ratings. Accessing a chapter is counted as one "read" and readers can "vote" (only positively) each chapter, in both cases contributing to the ranking of a story within the various lists in which it is included. Lists can be general categories – e.g. literary genres or topics – indexed by Wattpad on the homepage, but also created by readers. Within each list, stories can be

ordered by “hot” or “new,” but it is not clear how overall popularity or temporary bursts in readers’ votes affect the ranking, hence it can also depend on the timeframe of the reading activities, e.g. related to a new story update.

3.2.1 Identities and geographies

Some statistics about Wattpad readers can be obtained from the website itself: more than 90 million people use it every month, 90% of whom are “either Generation Z or Millennial,” i.e. younger than 35 years old (Wattpad 2018). More precisely, according to a survey conducted by Wattpad with English speaking users worldwide ($n = 650$), 80% of the users are female between 13 and 24 years old (Brady 2017). Among Generation Z U.S. Wattpad users, 58% identify as people of color, a few points percentage more than the national average (Wattpad 2018). The average daily reading time is 37 minutes, 85% of the times from a mobile phone (Wattpad n.d.; Miller 2015). The mentioned survey also states that “Almost half of respondents bought one or two books a month,” widely preferring paperbacks to ebooks, because readers “see themselves as curators and love to feel the paper, annotate their books, and financially support authors;” beside the fact that paper books have the advantage of not running out of battery (Brady 2017).

Fanfiction is the most published genre (more than 8 million stories, around 39% of the total), followed by Romance, Teen Fiction, and Action, all with more than one million stories (destinationtoast 2018b). Popular genres are already revealing of the identity of young readers: Teen Fiction – or Young Adult literature – is a genre whose frequent themes are friendship, “getting into troubles,” attraction for another person, and family issues (Wells 2003; cf. Cart 2016). Both the popular themes and the social affordance of DSR converge in creating a contact zone where teenagers can learn to become skilled social agents, which is one of the evolutionary functions of literature (Carroll 2004; B. Boyd 2009; Mar and Oatley 2008). Fiction is a simulated form of gossip, which allows us to learn about, and emotionally prepare for, possible social scenarios for which we would benefit to be ready (Dunbar 1997). Teenagers have normally little agency in society, but social media are spaces where they can express themselves as agents, while learning to navigate social relationships (boyd 2014; Milner 2004). DSR systems are an exponentially powerful tool to this end: they enable (i) the expression of agency typical of social media, (ii) the construction and negotiation of identities, roles, and social status with other teen readers, (iii) while reading about fictional teenagers often involved in difficult situations and relationships, (iv) with the possibility of discussing such issues with peers, (v) and with

the protection of the fictional distance, which leaves to them the freedom to choose to what extent to expose themselves as identifying with the characters or somehow relating to them.

Together with Simone Rebora and Gerhard Lauer, we decided to get to know more about Wattpad readers, their reading preferences, and how they respond to different kinds of stories and interact with other readers. Having a number of published stories estimated between 20 and 30 millions, one way to start exploring Wattpad is by looking at the stories' titles, which can be retrieved through the sitemap (cf. Pianzola, Rebora, and Lauer 2020 for more details on the methodology used and the interpretation of results). Using automatic language recognition tools (e.g. Sites 2013), it is then possible to estimate the extent of the linguistic and cultural diversity of the stories published. [Table 8](#) shows how many stories have been written for each language: English is by far the most represented language, amounting to around 79%. However, this is an overestimation, since English titles are quite often used for stories written in other languages, too. Overall, it is remarkable to find broad collections of stories written in non-Western languages. For instance, there are more Turkish stories than French ones, and more Vietnamese and Indonesian stories than German ones. Such a wide and diversified database of stories is an extremely useful resource for investigating the social dynamics of teenagers sharing a passion for reading and talking about their lives.

Table 8. List of the 13 most used languages on Wattpad (up to January 2018), identified with the support of the language detection algorithm CLD2

	Language	Number of titles	% of total
1	English	24,869,949	79.0%
2	Spanish	1,961,109	6.2%
3	Portuguese	684,004	2.2%
4	Turkish	560,919	1.8%
5	French	433,602	1.4%
6	Vietnamese	431,799	1.4%
7	Indonesian	304,601	1.0%

8	German	289,274	0.9%
9	Russian	264,352	0.8%
10	Italian	235,665	0.7%
11	Polish	213,150	0.7%
12	Malaysian	128,551	0.4%
13	Tagalog (Filipino)	118,331	0.4%

Beyond noticing a general cultural diversity, it is also possible to select a subset of stories and see the geographic location of readers who commented on them. From a small corpus of twelve English novels belonging either to the Teen Fiction or Classics genre, we extracted geographical location information for 35,000 users, and found that stories written in English are read by users living in many different countries. Figure 2 shows that most readers are from countries where English is either the first language (USA, UK, Canada, etc.) or a second language (India, Philippines, Indonesia, etc.), but there are also many non-English speaking countries, each of them amounting to around 0.5% of the readership (e.g., almost all European countries, Saudi Arabia, Egypt, South Korea, Argentina, etc.). Among the twelve stories in the corpus, Classic authors are all from the UK, and Teen Fiction authors come from the USA, Canada, and one from South Africa, but the works by all of them have been read in all continents. This means that there is a worldwide community of readers who read, comment, and discuss books, engaging with the reflections and opinions of people living in various countries, and meeting other culture-specific responses to the same stories.

Visit the web version of this article to view interactive content.

Figure 3: Map of readers' locations (n = 35,208) extracted from a corpus of twelve English language stories. Hover over countries to display the number of commentators.

Reader response is something that changes according to the reader's experiential background, their language, country of origin, and culture. By mapping Wattpad's users we can look at the cultural specificity of readers and see that our children and

students are talking about books with peers from all over the world. The extent of the mismatch between the language of texts and readers' location suggests that they could be native English-speaking migrants living in foreign countries, but also people reading in English as a second language. Filipino readership is particularly big: they not only write and read in their native language (cf. [Table 8](#)), they also read a lot in English, confirming that they are one of the most relevant groups among Wattpad readers (Wattpad 2019b).

Cultural studies have diversified and relativized the perspectives we take when we investigate reader response, contesting the dominance of white voices in literary studies (Said 1978; Spivak 1988). Wattpad shows that in the 21st century, thanks to digital technology, diversity in readers' response is also found simultaneously on the same reading platform. For both readers and researchers this is an opportunity to meet other perspectives on books, see how they affect people's lives, and start a dialogue about differences and similarities in how stories are received by readers with different cultures. The prosocial effect that fiction has (Dodell-Feder and Tamir 2018; Mumper and Gerrig 2017; Mar, Oatley, and Peterson 2009) is complemented by the opportunity of actual intercultural social interaction offered by DSR.

Another important aspect to underline is that Wattpad readers' construction of identity occurs exclusively through the text of their comments, unlike face to face reading groups, and it is affected by the public social context. Both the fact of taking place on a social-network platform and of participating in a reading group strengthen the performative aspect of identity construction (Peplow et al. 2016). The influence of social space on how readers present themselves is not something happening for DSR only, it has a historical antecedent in notes by famous readers. In fact, it has been noted that some of John Keats's annotations in the margins of books have "the signs of performance or at least of a consciousness of shared experience about it" (Jackson 2005). Accordingly,

in conversation with books as in conversation proper it is hard to say whether an opinion is being expressed in the conversation or formed by it. Readers annotating books must be forming their opinions *of the books* as they go, but the opinions about various subjects elicited by the text may be one or the other (preformed, ready-made; or made up on the spot) or a bit of both. (Jackson 2005, 136)

Wattpad comments in the margins offer an insight into the process of identity negotiation, reading comprehension, and the emergence/development of emotions and thoughts regarding a story.

3.2.2 Emotions and narrative interest

Wattpad data enables to dig deeper into books' reception by looking at the content of comments written in the margins of the story's paragraphs. One way to do this is by exploring the emotions expressed by readers, using computer-assisted sentiment analysis (B. Liu 2015) to see how the progression of a story elicit positive and negative emotions. By counting the number of positive and negative words, using as a reference a dictionary of words to which sentiment values have been preassigned (e.g. "love" = +1, "sad" = -1), Matthew Jockers (2014; Archer and Jockers 2016) and Andrew J. Reagan and colleagues (2016) have shown how to analyze the emotional arcs of stories. The same principle can be applied to look for patterns in the relationship between textual features and the emotional response they elicit. For instance, for *Pride and Prejudice* - the most popular text among the Classics, with more than 42,000 comments - comments have a more negative sentiment than the story, except at the beginning (Figure 4).

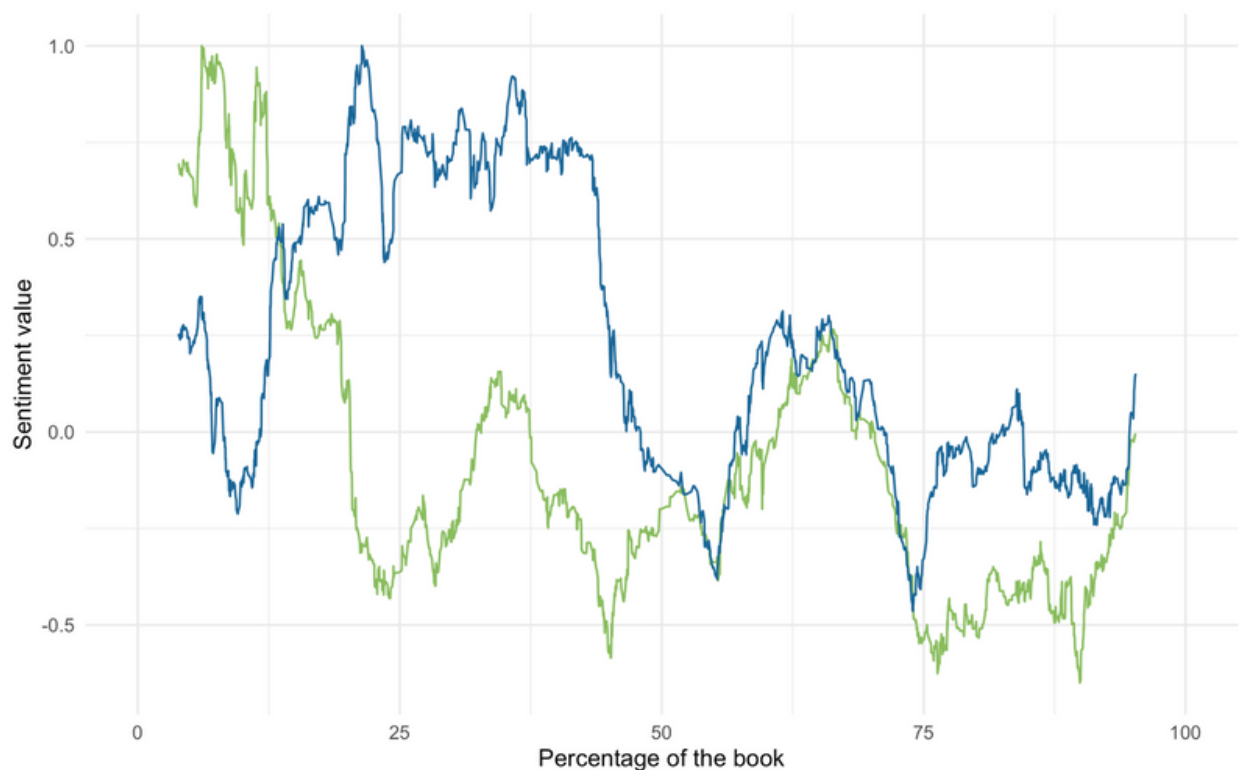


Figure 4: Emotional arcs for *Pride and Prejudice*'s text and comments

Looking at numerical data is not enough to understand the reasons for certain values, we have to actually read some of the comments to determine why they are so positive in the beginning. In this case, there are two plausible explanations, one related to the

story, the other related to contextual social habits. Browsing through the comments to the incipit of *Pride and Prejudice*, it becomes clear that its success on Wattpad is due to its mention in the Teen Fiction novel *After* (Todd 2013), in which one male character says: “Elizabeth Bennet needs to chill.” Many readers of *After* *enthusiastically* claim that they have been attracted to the Jane Austen’s novel for this reason: they liked *After* and express their positive expectations for a book mentioned in it. Indeed, Wattpad readers like to announce publicly that they started reading a story, often stating whether it is their first time reading it (“ftr”, first time reader) or they are rereading it (“rr”, rereading), and adding a motivation is a way to look for an affinity with other readers. The content-related motivation regards a dynamic between two characters: *readers impulsively dislike Mrs. Bennet – especially the way she talks about Elizabeth – and respond euphorically to Mr. Bennet’s wit and sarcasm when he talks to his wife in the initial scenes.*

If we extend this kind of sentiment analysis to a collection of books belonging to two different genres – e.g. Teen Fiction and Classics – we can see whether there are recurring patterns and whether they differ between the two genres. Statistical tests like correlation analysis and linear regression can reveal whether there are frequent associations between values of the story sentiment and values of the corresponding comments sentiment. What can be discovered is whether the emotional value of the story influences the emotions with which readers respond to it. That is, if a story uses many negative words related to anger and sadness, do readers verbalize their emotional response with more negative words than they do for positive stories? Figure 5 shows that there is a significant effect of the story sentiment on readers’ response and the two arcs have much more harmonious trends for Teen Fiction than for Classics. This result is not trivial because the verbalization of emotional response to a story can take many forms and it is not obvious that positive words in stories trigger readers’ positive utterances. Values for Teen Fiction are quite reliable, being based on hundreds of comments for every paragraph of the story. On the other hand, Classics often only have a handful of comments and this scarcity affects the computation of sentiment scores. To overcome this limitation, we can explore more in detail the paragraphs and comments in which there are peaks in the sentiment values or discrepancies between the two emotional arcs.



are reported. This kind of positive reaction is often elicited by a contrast between a candid statement and a negative character that previously acted or talked in an offensive or annoying way. For instance, between Mrs. Bennet and Mr. Collins, in *Pride and Prejudice*, or the various “bad boys” and “queen bees” in Teen Fiction novels. Witty characters are very much appreciated by teenagers and having secondary characters whose personality create a contrast is a strategy that amplifies readers’ reaction. For example, Miss Bingley with respect to Elizabeth, and the double contrast between the proud but warm-hearted Mr. Darcy and the charming but manipulative Mr. Wickham. A similar process can also work for a single character, like in the case of Tessa – the protagonist and first-person narrator of *The Bad Boy’s Girl* – who lives a continuous internal conflict: first, between the attraction for Cole and her resistance due to childhood memories; second, between the love she feels for Cole and the will to keep him at a distance after he hurt her. Comments show that readers mimic a similar conflict, especially after an extreme inversion of sentiment in the story (chapter 32, 75-80% in Figure 5: they feel for Tessa and hate Cole because he made her suffer, but they also hate Tessa for not forgiving him, and empathize with Cole).

The relevance of this pattern is confirmed also by looking at the most commented paragraphs. Peaks in the number of comments are related to story events that have a strong narrative interest, i.e. related to suspense, curiosity, or surprise effects (Sternberg 1992), not necessarily related to emotionally intense scenes. For instance, reading the anticipatory chapter’s title “What It Feels Like To Get Your Heart Broken” (*The Bad Boy’s Girl*), many readers are emotionally expressing their concern for what they are about to read. This is only partly in line with one of the main functions of highlighting e-books, where users usually highlight “affective climaxes of a narrative” (Rowberry 2016). Here, reader response is not strictly linked to the story content but rather to its rhetorical and narrative organization. Moreover, in highly commented paragraphs on Wattpad there is no trace of the commonplaces identified by Rowberry, which are characterized by “perceived wisdom:” moral values do not spark many discussions among Wattpad readers. This might be due to genre-specific limitations of Teen Fiction and Classics, or to the readers’ young age, but highlighting and commenting also seem to have two distinct functions, even though both activities are publicly shared and not meant for individual use only. Similarly, more prestigious books have fewer highlights and notes than popular and free books, also prompting different kinds of behaviors: the former have notes reflecting attentive and critical reading, the latter have a lot of chatting (Barnett 2014).

Another remarkable aspect is the great number of comments on initial paragraphs, which is an implicit sign of emotional engagement, since some readers express their excitement for the beginning of a new reading experience, while others claim that they are rereading the story, plausibly because they felt involved during their first reading. The latter sometimes also write comments addressed to other readers, warning them when there may be spoilers. Indeed, interactions between readers are a significant part of comments on Wattpad.

3.2.3 Interactions and community

When reading Teen Fiction and Classics, Wattpad readers interact with each other in different ways: for Teen Fiction, 22% of the comments are answers to other commentators (secondary thematization; cf. section 2.2), while the percentage increases to 31% for the Classics. To better understand how they interact with the text and with each other, we can use network analysis.

Figure 6 shows the commentators and chapters network graph for *Pride and Prejudice*. The first chapters are isolated and clearly surrounded by many commentators who do not have connections to other chapters. This means that they probably did not continue reading past the beginning of the story, as it is also confirmed by the higher number of comments to these chapters in comparison to the others. The greater number of pink edges and the clustering of pink nodes show that readers talk between each other more than directly commenting on the text. Comments to the first chapters are often questions like “Is this a complete copy?” or “i dont understand... is this the real book?,” but also “Is this for real? Jane Austen has a wattpad account... Do you have a copy right or whatever? Just asking..”. These kinds of questions are likely to start conversations between users before they actually begin to engage with the story; they are annotations with a social function (cf section 2.3; Thoms and Poole 2017). The cluster on the right is stretched because of the attraction of a very sociable commentator (USER_277), who had many social interactions from the beginning of the book and, thus, pulls the edges connecting the clusters towards the top.

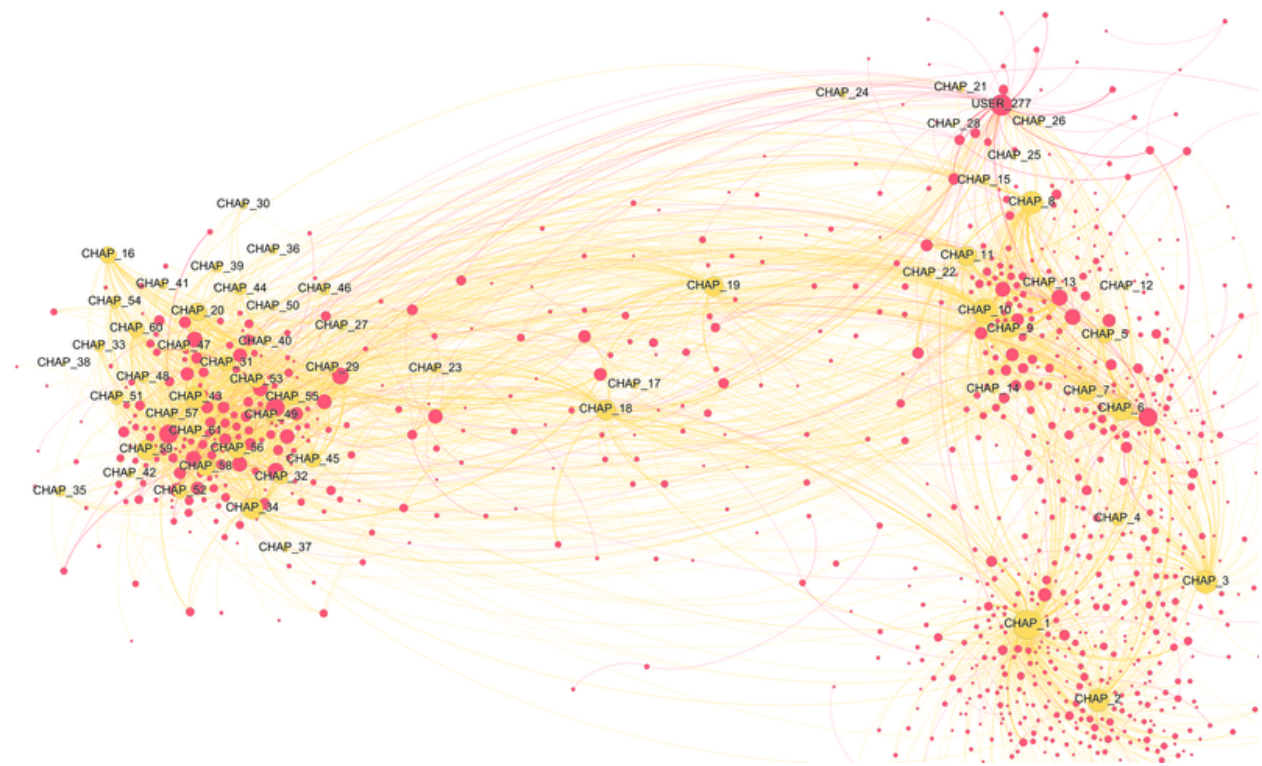


Figure 6. Commentators and chapters network graph for *Pride and Prejudice* (detail).

Commentators and their interactions are in pink, chapters are in yellow. The full graph contains 1,067 nodes, 17% of the total in the book (6,278). This reduction was obtained by taking into consideration only nodes that interacted between each other at least 3 times. Weighted degree of the nodes (i.e. their importance in the network) is visualized through a logarithmic scale, thus half diameter corresponds to 1/10 of the weighted degree. This option allows to better highlight distinctions.

To better understand the motivation of the higher sociability of *Pride and Prejudice*'s commentators we have to read the comments. The resulting insight is that, since Classics are written in an English that is difficult to understand for many teenagers, and depict worlds with social norms not familiar to them, conversations in the margins are often requests for help to understand what is happening in the story. This phenomenon suggests that the resistance of young readers to reading novels written in an English difficult to understand can be overcome by peer learning and collective intelligence (S. Evans et al. 2017; Guilmette 2007; Jenkins 2006), since readers explain each other paragraphs that they did not understand. In this regard, one of the most remarkable examples is that of a commentator who regularly paraphrased the most difficult paragraphs, receiving the praise and thanks of many other readers. A similar form of collective intelligence also emerges from interactions on fanfiction websites,

where readers collaborate in figuring out plot developments but also details about the story setting (Thomas 2011b), and also in DSR learning activities in smaller groups (Thoms and Poole 2018).

The Bad Boy's Girl graph (Figure 7) shows that there are many commentators who are very active, as attested by the size of the pink nodes, but also a few single commentators who form separate clusters. Clusters of readers are so relevant in the network that they can even attract chapters, isolating them from the main cluster, as in the case of Chapter 5. Its position is not due to low engagement (it ranks 32nd out of 42, with 46,700 comments vs. 16,900 of the least commented chapters), but rather to the attraction of a reader who commented it a lot and had another user as a privileged discussant. In general, there are more user-user interactions for Classics, but readers of Teen Fiction generate groups with stronger bonds, that is readers that are able to interact with other readers more than 250 times (USER_682).

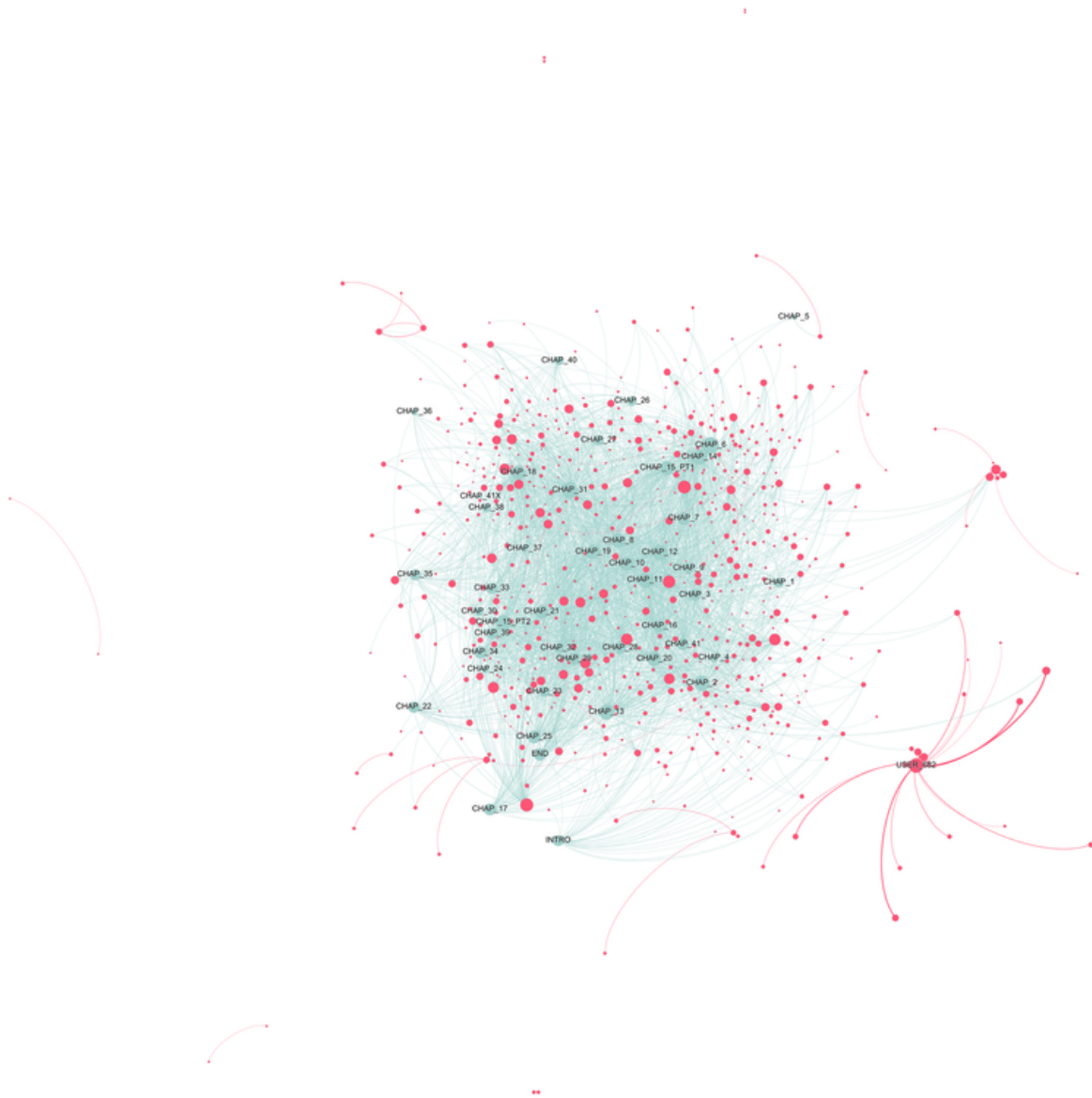


Figure 7: Commentators and chapters network graph for *The Bad Boy's Girl* (detail; Pianzola et al. 2020).

Commentators and their interactions are in pink, chapters are in turquoise. The full graph contains 735 nodes, 0.6% of the total in the book (130,615). This reduction was obtained by taking into consideration only nodes that interacted between each other at least 30 times. Weighted degree of the nodes (i.e. their importance in the network) is visualized through a logarithmic scale, thus half diameter corresponds to 1/10 of the weighted degree. This option allows to better highlight distinctions.

By applying different techniques to study comments in the margin, we found that teenage readers love witty characters, conflicts of affects and values, and cultural references that are familiar to them (another of the themes in the most commented paragraphs). But we need to be careful in generalizing about readers' behavior, since the texts and the readers on Wattpad are only a subgroup of all kinds of literary texts and readers. The young age of Wattpad readers and their still relative small cultural capital may affect the way they talk about books, as it also happens on other kinds of digital platforms (Rehberg Sedo 2011b; Thomas and Round 2016; Thomas, Bray, and Gibbons 2015). Analyzing the way in which Wattpad readers engage with stories and between them, it is possible to identify patterns related to narrative interest, aesthetic values, and social interaction. When reading Teen Fiction, social-bonding (affective interaction) is prevalent, when reading Classics social-cognitive interaction (collective intelligence) is prevalent. From an educational perspective, it is relevant that readers who uses Wattpad learn to read Classics and to judge books not only in direct emotional response to characters' behavior, but focusing more on contextualized interpretations of the text. And the path to such important achievement is strongly supported by the kind of peer-to-peer interaction enabled by digital social reading.

Overall, by examining reader response at scale we can set the ground for statistically valid claims about how many young people read. Namely, we can understand what kind of emotions readers feel towards the characters and how they react to specific narrative strategies like suspense. Ultimately, comments are a resource for different kinds of inquiries, investigating how, while reading, values are negotiated, identities are constructed, authors interact with readers, metaphors are interpreted, intertextual connections are built, factors like experiential background (Caracciolo 2014) and personal relevance (Kuzmičová and Bálint 2018) affect reader response, etc. It has been debated whether marginalia can actually help to understand reading practices of the past (Jackson 2001; Kingten 1996; Rose 1995; Martin 1994; Machor 1992; Darnton 1995). In this regard, studying Wattpad comments to understand contemporary reading practices has the advantage that the retrieved information can be compared with other kinds of data within more comprehensive reading research programs (Rebora and Pianzola 2018). For instance, the quantitative analysis of comments can be used to identify readers that commented a lot and invite them for interviews about their reading experience.

3.3 Fanfiction

Table 9. AO3 features

DSR platform	Content macro-category	Relation to source text	Type of audience	Timeframe
AO3	Comment in footer	Immediate	Institution, global	After reading, self-paced, scheduled
	Rating	Immediate, external	Institution, global	After reading, self-paced
	Review	External	Institution, global	After reading, self-paced
	Tag	External	Institution, global	After reading, self-paced
	List	External	Institution, global	After reading, self-paced

On AO3, comments can only be written in the page footer. On top of the page readers can see the story metadata, including its rating: “hits,” i.e. number of times a story has been accessed, and “kudos,” i.e. positive reactions to the story. Ratings are also displayed at the end of the story, before the comment section. Tags, in the sense of labels assigned by readers, can be used when bookmarking a story. Bookmarks automatically include all the tags the work's creator used, but readers can add their own, even supplementing them with notes, which can be used to write a review, a description of some of the story features, or a summary. “Rec,” i.e. recommendation, is a specific bookmark that can also be used as a filter to compile a list of recommended stories, therefore also working as a form of rating. The type of audience for all sorts of reader-generated content is tied to the restrictions chosen by the author – to make the story visible either to anyone or only to AO3 registered users – whereas readers can only choose to make bookmarks private. This combination of features makes AO3 an example of “Manifold discussion within a community” (DSR type 2).

3.3.1 Fanfiction as reader response: diversity

Fanfiction is an emblematic manifestation of creative and critical reading (Jenkins 2013). It is often mainly associated with the creation of derivative work – i.e. text written by fans of a narrative universe created by someone else – but its being a form of reading is the precondition for being a derivative work of fiction. Participation in

fanfiction-related activities is not only about writing stories, it is primarily about reading, reviewing, and discussing with other fans; it does not necessarily lead to authoring (Magnifico, Lammers, and Curwood 2020). And even when it is about writing a new story, fanfiction is content created in response to reading a source text, before being itself a source text subject to further reader response. As such, even when it becomes the topic of primary thematization by fanfiction readers, it is also almost always discussed from a perspective of secondary thematization, i.e. discussing a comment on the fandom's source text(s).

In 2013, four years after the AO3 opening (November 2009), a fan started an online census, in English, of AO3 users (n = 10,005; Lulu 2013). The mean age of the respondents was 25, 77% between 19 and 29 years old and 16% between 16-18 years old. With respect to gender, 80% identified as female, 4% as male, and the others chose options related to a more fluid gender identity. Only 38% identified as heterosexual. With respect to "ethnicity," 76% reported to be White, 2% Black, 7% Hispanic, and 5% Mixed. To sum up, three quarters of AO3 readership is composed of White non-heterosexual women in their twenties. Frequency of usage was quite high, with 49% of respondents reporting to use the site at least once a day, and 25% several times a week. The median value for usage length was in the range of 1-2 hours per day, 5.6 hours per week. The census disproved the widespread opinion that all fanfiction readers are also authors themselves – only 38% reported to post stories on AO3, although according to a different survey 58% are also authors (n = 5,528; Organization for Transformative Works 2013) – and confirmed the importance of reading socially: 74% of respondents left kudos on stories (93% for OTW survey), 44% (82%) commented on them, and 56% (52%) bookmarked them.

Most AO3 stories are read on average by 455 readers (median of hits per chapter; n = 1,118). As a comparison, most Wattpad stories are read on average by 20 readers (n = 5,295) and those in its fanfiction category by 44 readers (n = 685; destinationtoast 2019b). In total, there are more than 6.5 million works on AO3 (October 2020), making it probably the biggest fanfiction archive for English-language stories (more than 90% of total). It certainly is the fastest growing one (fffinnagain 2017) and the DSR platform with the most active readership, gathering every month more than 16 million kudos and two million comments (AO3 Admin 2020). AO3 readers are quite omnivorous, having a wide variety of genre interests: Romance is read by 79%, Fluff (stories whose focus is the display of affection) and Smut (major explicit sexual content) are read by 74%, while Angst (stories aiming at eliciting a feeling of unrest and uncertainty in readers) and Humor are read by more than 60% (Lulu 2013).

Moreover, there is a marked preference for long narrative (70%) rather than short stories (43%). The great majority of readers (90%) likes to read about male homosexual relationships (“M/M”), but half of the total also reads stories with no romantic or sexual content (“Gen”), and only 50% reads about heterosexual relationship (“F/M”), suggesting that many fanfiction readers are probably unsatisfied with the heteronormative catalogue of many publishers and entertainment companies.

Further insight about diversity comes from Milli and Bamman’s work (2016) on Fanfiction.net and their finding that characters which are less prominent in the canonical texts receive much more attention than central characters in fanfiction (n = 1,000). There could be two explanations for fans’ preference for secondary characters and unexpected relationships: one related to the under-specification of the character’s personality and life story, which leave more room for the fans’ imaginary explorations, the other reason is related to the under-representation of gender, racial, and ethnic identities in prominent roles in mainstream literature, which leads fans to self-identify with such minor characters and to be willing to tell their stories. However, the preference for F/M and M/M fiction, suggests that readers’ imagination likely draws on hints found in the canon, where most major characters are men (‘The Geena Benchmark Report: 2007-2017’ 2019; H. Anderson and Daniels 2016).

Another possibility is that women use male characters and relationships as a proxy to explore female desire free from the oppression they live in everyday life (Busse and Lothian 2017). What motivations are behind these preferences is uncertain, as shown, for instance, by two trends in the recent *Star Wars* fandom: the wide preference for the White queer pairing Kylo Ren/General Hux (#kylux, antagonist and minor character) – to the expenses of the popularity of the colored queer couple Finn/Poe (#stormpilot, two main characters) – and for the white heterosexual Rey/Kylo Ren (#reylo), cutting out the Black Finn, whose plot trajectory is closer to that of Rey (Pande 2018; Stitch 2019). However, the situation is more complicated, because the data show that #stormpilot had an initial peak similar to #reylo but then drastically declined in popularity (destinationtoast 2016; cf. Messina 2019; and Lulu 2020 for data about other fandoms).

As an example, let’s consider Harry Potter fanfiction, one of the largest on AO3 (245,230 stories in October 2020, excluding crossovers with other fandoms). Looking at the tags that authors add when posting a story online we can gain a lot of information about diversity. With respect to sexual orientation, there is a big difference between canonical literature and fanfiction. In the Harry Potter canonical novels, only

Albus Dumbledore and Gellert Grindelwald are homosexual, whereas 61% of fanfiction stories has a queer pairing (a value in line with the overall AO3 average), that is nearly 150,000 stories. This is not the case for all fanfiction websites, e.g. Fanfiction.net only has around 20% of queer relationships overall, and they amount to around 22% in Wattpad fan works (destinationtoast 2019a; franzeska 2018).

A factor that could influence the preference for certain characters or relationship is related to the attention given to female and male characters in the canon. I already mentioned that, in general, women get less screen and dialogue time than men ('The Geena Benchmark Report: 2007-2017' 2019; H. Anderson and Daniels 2016; Kagan, Chesney, and Fire 2020), but looking at individual movies we can assess that female characters are also less likely to have a big part of all the dialogue in a movie (destinationtoast 2018a). This is relevant because the amount of time a character talks in a movie is correlated to the attention that the same character gets in fanfiction ($R^2 = .30$). Thus, it could be an important factor in influencing a character's popularity on AO3. An analysis of data from Fanfiction.net showed that, on average, 40% of character mentions in the canon are of women, but in fanfiction this ratio increases to 42% ($p < .001$), suggesting that fanfiction does indeed devote more attention to female characters (Milli and Bamman 2016).

Let's see what happens in *Harry Potter* fanfiction. I decided not to compare fanfiction to the original novels because they are copyrighted, but I found data about characters' appearance in the movie series, so I will compare screen time to the characters' tags found in AO3 stories, which indicate the presence of a character in a story. The two metrics do not really have the same meaning but are a good approximation of the attention a character receives from institutional authors/producers or from fans. Moreover, a lot of the variation in fans' attention is associated to the variation of character's screen time ($R^2 = .91$, *Pearson* $r = .82$). However, there are only eight movies but thousands of fanfiction stories ($n = 160,453$), hence it is more likely that minor characters receive more attention in fanfiction than in movies. For instance, it is possible that a character appears in all fanfiction stories, having a presence of 100%, but no character is on screen the whole time. Nevertheless, the number of characters I considered is the same and I looked at relative proportions within each medium. Keeping these differences in mind, I tested (i) whether female characters are represented more or less than male characters in fanfiction, and (ii) whether the gender gap is bigger or smaller in comparison to the movies.

Among the 122 characters appearing in the movies, 46 are female and 76 are male: a 65% gender gap, i.e. 5 men every 3 women. With such initial disparity in the gender of the characters that fans can write about, I would expect an exponentially wider gender gap in terms of characters' occurrences in AO3 stories, merely for a probabilistic reason: the more characters there are of a certain gender, the more likely it is that they will be part of a story. This is indeed the case, there is a wider gap in the number of characters of each gender that are featured in fanfiction stories in comparison to movies. The answer to the first question is that female characters are represented far less than male characters, being featured 2.5 less times in stories. But this gap is difficult to interpret in itself: does fanfiction do better than canon movies in terms of gender equality or it simply reflects an underrepresentation of women occurring in movies in the first place? To have a better idea we can compare the characters' occurrences in stories to the screen time they have in movies. Indeed, female characters have 3.2 times less screen time than male characters, meaning that on average fanfiction reduces the gender gap in characters' representation by 34%.

In Figure 8, I plotted how much characters are represented as a percentage of the total screen time and total *Harry Potter* AO3 stories. The dashed line indicates an even ratio between screen time and fanfiction occurrences, thus characters positioned above it received more attention in fanfiction than in movies, whereas the opposite is true for characters below the line. Remember that, because screen time and AO3 occurrences measure characters' representation in two different ways, absolute values in increase or decrease should be interpreted carefully. It is rather the importance of characters relative to each other that should be observed: for instance, in fanfiction stories professors Slughorn and Umbridge are much less popular than the students Blaise Zabini and Billy Weasley, although the former have more screen time than the latter.



Figure 8: Proportional representation of characters in the *Harry Potter* movie series and in AO3 fanfiction (detail).

The percentage beside the characters' names indicates the increase or decrease in representation, calculated as a percentage variation from screen time (seconds) to occurrences in AO3 stories. Axis have been cut to show trends for most of the characters.

There is also an additional consideration to be made. As I already said, the fact that the majority of main characters are male increases the overall probability that men become more popular than women, and the proportional weight of very popular characters can possibly hide a variation of the gender gap in the representation of minor characters. Indeed, if we look at the median, AO3 reduces the gender gap by 46%. This is equivalent to saying that, most female characters appear on the screen

1.7 less time than most male characters, but on AO3 the ratio is reduced to 1.4 less occurrences for women in fanfiction stories. The gender gap is still there, but it is likely due to an imbalance in the initial number of male and female characters, and to their uneven presence on screen. Among the ten characters with most screen time, only three are women.

If we exclude the most popular characters and the extremely minor ones, we can see a clear trend: the majority of female characters has an increase in representation, whereas the majority of male characters has a decrease (Figure 9). However, using a multilevel mixed effect model for repeated measures, I found that the reduction of the gender gap from movies to AO3 stories is not statistically significant (for the model with no outliers, $b = -0.018$ [bootstrapped CI $-0.408, 0.341$]). This does not mean that fanfiction has no effect at all on the gender gap in representation, rather there are other causes that influence the popularity of characters more strongly than gender. Some of them may concern canonical characters' personality and plot, some other reasons may be related to context and biases that are specific to the platform, like AO3 readers' preference for M/M stories, which is different from Fanfiction.net and Wattpad (destinationtoast 2019a).

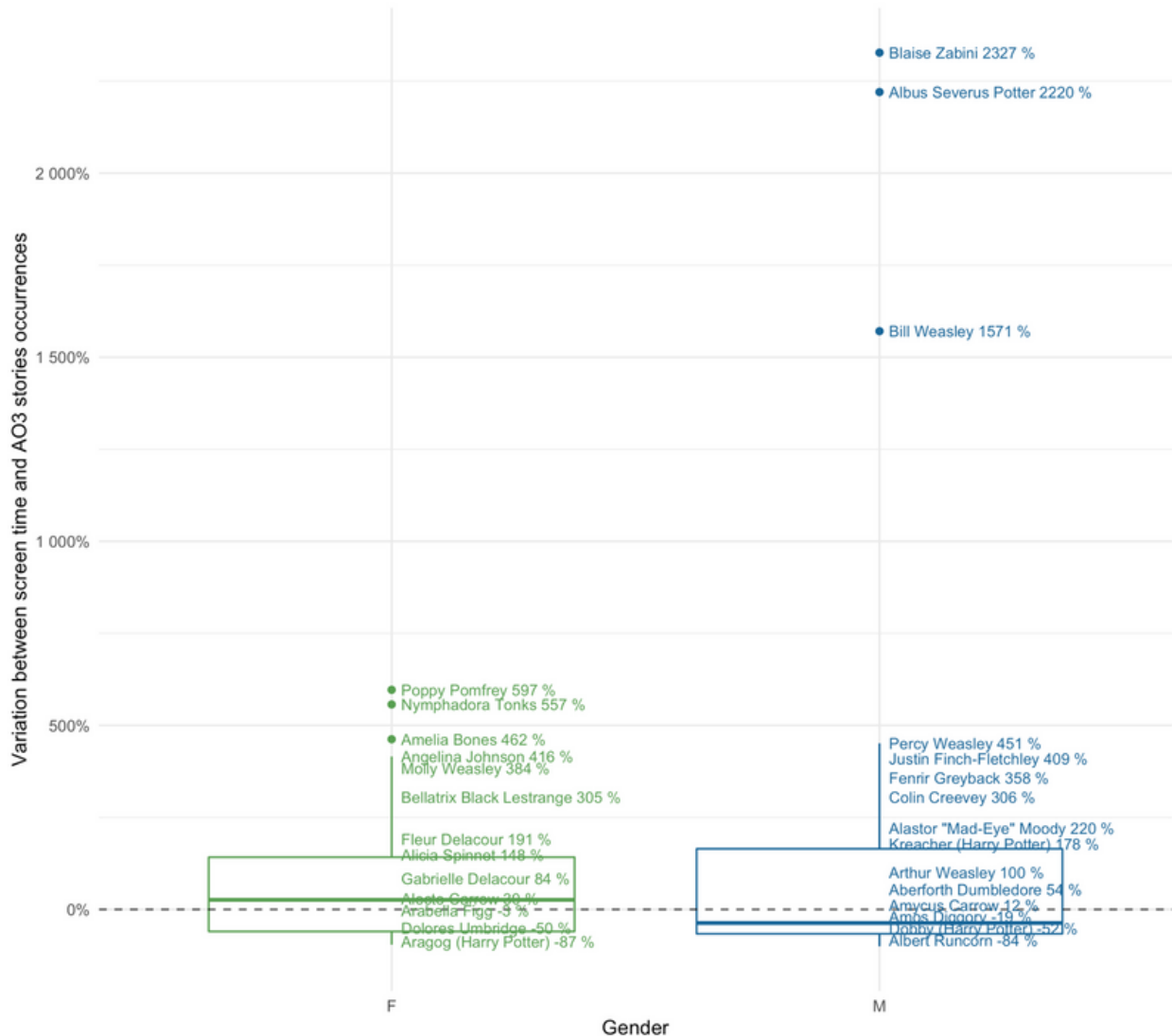


Figure 9: Variation between screen time and AO3 stories occurrences for the majority of male and female *Harry Potter* characters (n = 100). Outliers have been excluded for both screen time and AO3 occurrences. Only the names of some characters are shown, as an example.

Figures 8 and 9 show that the Black student Blaise Zabini has a huge boost in popularity, suggesting that maybe racial equality is another improvement undertaken by fan readers (but remember that, for legibility, I excluded from these figures the most popular characters, who have even bigger boosts). To better understand racial diversity on AO3 we need to investigate further.

I previously mentioned that 24% of AO3 readers identify as non-White, however only around 9% of *Harry Potter* fanfiction stories has a Person of Color tagged as character, and 0.2% features exclusively People of Color. In comparison, the canon is even less

diverse, since in the eight movies of the *Harry Potter* series there are 12 people of color (10%), who in total speak for 0.5% of the screen time (Marron 2015). The trend is similar for the whole publishing industry. For instance, if I want to read Young Adult fiction written by Authors of Color and I am living in the UK, by entering in any bookshop – physical or online – I can only choose around 8% of the titles published in this category since 2006 (Ramdarshan Bold 2018; cf. So 2020 for US data). The percentage is the same as for AO3, so it may seem that there is not much difference between institutionalized literature and fanfiction, but 8% of the UK 2006-2018 catalogue means that I can choose among 600 books, whereas on AO3, considering all fandoms, I can choose among 11,000 stories explicitly tagged “Character(s) of Color.” A closer look at *Harry Potter* fanfiction reveals that there are many more relevant stories but they are not explicitly identified by this general tag: 8% of Harry Potter fanfiction corresponds to 18,000 stories with at least a Character of Color, plus 366 stories with no White characters at all. Many more than 11,000 stories across all fandoms. My guess is that the tag “Character(s) of Color” is used for stories where such characters have a major role. As a comparison, on Wattpad I found at least 1,500 stories tagged “poc” (People of Color) published in the same time span. Thus, fanfiction seems to proportionally mirror the little racial diversity found in institutional fiction, but in absolute numbers it offers many more stories to readers.

More generally, we can look at how many different ways fanfiction authors have to creatively reinterpret the characters of the stories they like. One way to do it is through the “relationships tags” used on AO3 – which refer to family, romantic, or erotic relations – and the “additional or freeform tags” – which are used to specify various aspects of a story, including themes or peculiarities a character has. In this way, we see that in 2002 fans were writing stories about 15 different relationships in which the character Harry Potter was involved, in 2010 there were already 75 different possible combinations, and in 2019 Harry was participating in 820 different relationships, including those with characters from other fandoms. Other characters followed a similar pattern of increasing diversification, e.g. Draco Malfoy went from 2 to 14 to 86, and Hermione Granger from 1 to 6 to 80. More creativity is required to explore new aspects of a character’s personality, nevertheless over the years there has been an increasing research for original and more nuanced characters’ representations (Figure 10).

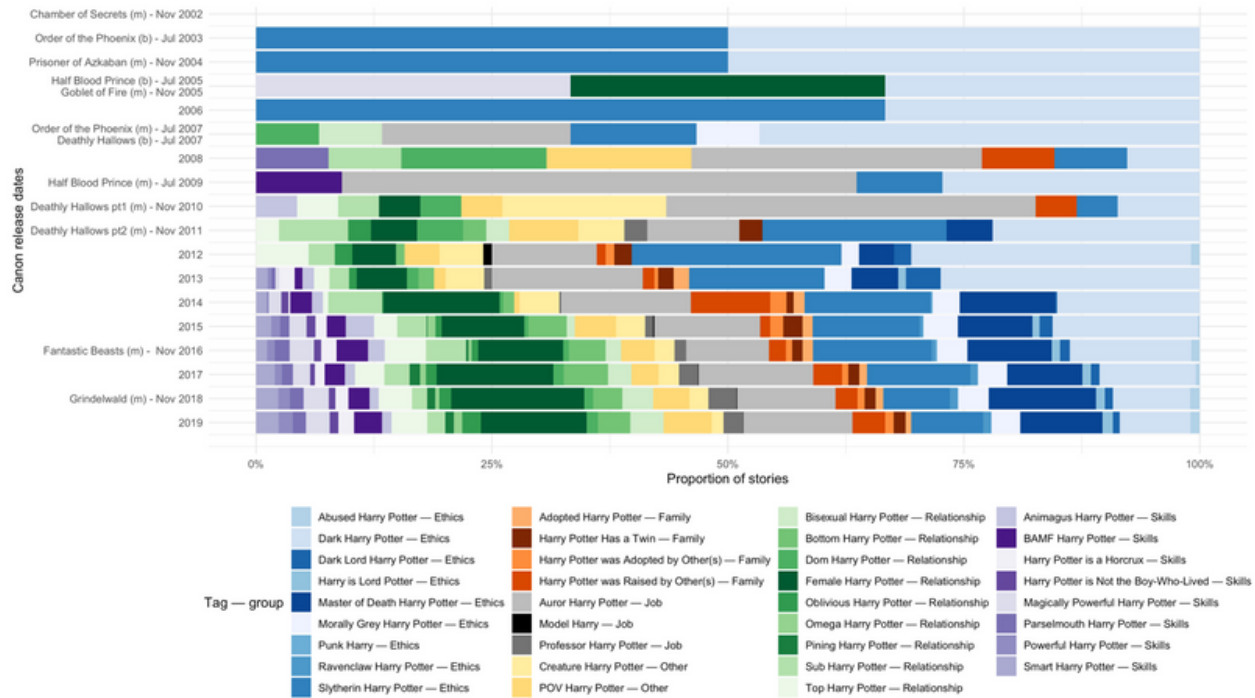


Figure 10: Proportion of stories with some specific character tags in relation to all AO3 stories featuring the character Harry Potter. Books (b) and movies (m) are listed beside the year in which they have been published/released.

Over sixteen years, fan explored increasingly more ways to characterize Harry Potter, in some cases picking hints of his personality as it is presented in the canonical novels or movies and developing them further, in some other cases exploring fanfiction tropes related to sexual orientation and relationships. Strongly reliable data for AO3 start from 2010, following its opening in November 2009; data for previous years are for stories imported years after they have been written, so they are not representative of all AO3 users. Anyway, the main takeaway is that readers' interest is attracted by three macro areas (categorized by me, not intrinsic to AO3 tagging system): one related to ethics, one related to Harry's ties with other characters (Family and Relationships), and one related to the abilities he has and puts into action (Skills and Job). If you are curious about the meaning of specific tags, I suggest consulting the fans' encyclopedia Fanlore.org. What I want to highlight here is that, over the years, new specific interests have emerged or have become more popular, in part because of the influence of newly released movies. More specifically, a steadily growing topic is that of relationships, namely of the kind related to the "BDSM" and "Omegaverse" alternate universes, two fanfiction tropes that became popular across many fandoms in the last ten years ('BDSM AU' n.d.; 'Alpha/Beta/Omega', n.d.; Busse 2013).

Being a set of conventions not directly related to any specific fandom, the BDSMverse and the Omegaverse show how fans construct their own strong identity as readers who have a preference for certain themes – or need to cope with them through the mediation of fiction – and share hermeneutic norms. The scarcity of such themes in mainstream literature brought to the emergence of practices of reading-writing (*lectoescritura*) as a form of reader response and affirmation of the rights of readers over that of authors and publishers. Indeed, AO3 is an explicitly noncommercial and independent space, standing out from both the majority of web platforms and the traditional publishing industry. Henry Jenkins said that “Fan fiction is a way of the culture repairing the damage done in a system where contemporary myths are owned by corporations instead of owned by the folk” (Harmon 1997). In this sense, AO3 is one of the most coherent expressions of this cultural phenomenon, where stories are written, shared, and read devoid of any commercial interest or authorial worship.

Its code is open source, and the archive has been designed, coded, and maintained nearly entirely by the community it serves—a community made up mostly of women. Because the controversy that sparked its existence was surrounding a disconnect with the community’s value system, baking these values into the design of the site was a priority. As a result, the design of AO3 is a unique example of building complex values and social norms into technology design. (Fiesler, Morrison, and Bruckman 2016).

If principles of the attention economy (Davenport and Beck 2001) still apply to such a non-commercial space, they are not related to monetization or the website infrastructure and design (Minkel 2020) but to psychological reward and the negotiation of reputation within the community (cf. Tushnet 2007; Meng and Wu 2013). Accordingly, one of the most interesting aspects of fanfiction in terms of cultural history is that, on the one hand, it is similar to non-urban societies where stories are not commodified; on the other hand, it is the existence of leisure, enabled by industrial modernization and organization of labor (Burke 1995), that guarantees the condition of possibility of fanfiction, which is created by amateurs during their time off work.

Before moving to analyze the other side of fanfiction, that is the response fan authors get from fellow readers, I have to mention one last thing: “meta.” “Meta” is a term (and a tag) used to refer to all kinds of non-fiction about a canonical or fan work, usually in the form of an essay, a short opinion piece, or, more recently, even data analyses (‘Meta’ n.d.). On AO3 there are currently around 10,000 works tagged as

“meta,” around 400 only for *Harry Potter*. Meta works are an additional type of content – similar to social media post – that qualifies AO3 as an example of “Manifold discussion within a community.” I will not go into more details here, but within the fanfiction universe they can be considered the most articulated form of amateur literary criticism.

3.3.2 Reader response to fanfiction: peer learning

Besides being readers who respond to canonical fiction, fans are also authors whose work become subject to other readers’ response. From the perspective of DSR, this is a shift of focus of primary thematization, from commentary on the fandom’s source text(s) to commentary on the creative work of fans, although reference to the canon is often part of comments, thus making comments on fan works also a form of secondary thematization. Since readers write feedback directly to the authors – not discussing something written by a third person – there is a lot of emotion work involved (Hochschild 1983) in comparison to, for instance, book reviews (Reagle 2015). When commenting on other fans’ works, fans consider how these authors may feel when receiving feedback, because often the commentators themselves have received feedback on their work and reacted emotionally to it (Campbell et al. 2016). Such emphasis on emotion work tells fans’ commentary apart from other kinds of amateur literary criticism, making it easier to create an inclusive and community-like atmosphere on fanfiction platforms.

The Organization for Transformative Work, the institution that created and now manages AO3, surveyed fans about their perception of AO3 (n = 3,746). Among other questions, they asked “How would you, personally, describe the Archive's atmosphere?” and, rather than giving a range of closed options, they allowed fans to reply to the open question, then manually tagged the answers and did a word frequency analysis (Organization for Transformative Works 2013). The most frequent terms used were “friendly” and “welcoming,” with “community” and “inclusive” being also among the top words. More specifically, manual tagging revealed the following positive perception: “friendly/supportive” (25%), “welcoming/homey” (9%), “open” (8%), “comfortable/safe” (5%), “inviting/appealing” (3%), “pleasant/nice” (3%), “community-like” (2%), “inclusive/accepting” (2%), “diverse” (1%). On the other hand, the most frequent negative answers were “lacking community/empty” (4%) and “impersonal” (4%). The general impression seems to be positive, thus it would be interesting to explore what kind of comments create such a positive atmosphere and whether they have an effect on the storytelling skills of fans. Cecilia Aragon and

colleagues tested it in different ways, developing a theory of “distributed mentoring” (Aragon, Davis, and Fiesler 2019; Frens et al. 2018; S. Evans et al. 2017; Campbell et al. 2016).

Drawing on Phillip Dawson’s (2014) mentoring framework, Aragon and colleagues define social relationships on fanfiction sites as a form of mentoring whose interactions are informal, weakly tied, peer-based, many-to-one or many-to-many, and occurring exclusively mediated by digital technology. In this context, fan authors improve their writing skills thanks to the feedback received in the various comments left under their stories, but also engaging in conversations with some of the commentators. This phenomenon has been identified thanks to ethnographic qualitative research (S. Evans et al. 2017; Campbell et al. 2016), but it has also been investigated with large-scale quantitative methods on Fanfiction.net. The results show that the number of reviews received by an author has a positive effect on the lexical diversity (MTLD) of the chapters that the same author writes in the following months (16.6 million chapters, 1 million authors; $p < .001$, Cohen’s $f^2 \approx 0$). However, given the extremely small effect size of this statistical relation, we have to consider the serious limitation of this methodology based on measuring lexical diversity, as also acknowledged by the authors of the research (Frens et al. 2018).

Beside an increase in lexical richness in time, with Alberto Acerbi and Simone Rebora (2020), we observed a more evident increase in the complexity of the stories written in the last 18 years and published on AO3. Overall, the average number of characters featured in a story doubled, from three in 2002 to six in 2019; and the number of relationships that fan authors are able to plot increased from 1.2 to two over the same time span. These results suggest that fanfiction authors became more skilled in telling stories involving many characters and their various relationships.

Aragon and colleagues focused only on the process of becoming better writers, so the term “distributed mentoring” is appropriate, but I prefer to use the term “peer learning,” because the feedback received by fanfiction authors can also help them to improve skills needed for critical reading, without necessarily aiming at crafting something themselves. In a context in which intertextual reference to the canon is crucial, more experienced readers can point out problematic interpretations or errors, increasing awareness and suggesting new ways of reading. For instance:

Okay, I have one slight issue. Hermione doesn’t learn the Patronus charm until her fifth year, so did being inside of Harry’s head give her the ability to do it? If so,

how would she know it was an otter if she wasn't supposed to have learned it yet.
(Campbell et al. 2016, 7)

Comments of this kind are helpful not only for authors, but for other readers too, who can easily find them below the stories they read.

To sum up, thanks to a their diversity of cultures and topics, Wattpad and fanfiction platforms seem to be places where people are able to find “intimate serendipity, a place where they can express an authentic sense of self without fear of attack, manipulation, or unusual exposure while remaining open to things that will surprise and delight them” (Reagle 2015). This is not the case for DSR platforms whose primary purpose is book evaluation to support the management of information and purchase decision, like in the case of book reviews websites. However, platforms like Goodreads, where readers can elaborate on their reads and organize books in different lists, can help us understand other aspects of reader response and popular uses of concepts related to literary genres.

3.4 Cataloguing, rating, reviewing

Table 10. Goodreads features

DSR platform	Content macro-category	Relation to source text	Type of audience	Timeframe
Goodreads	Highlight	Immediate	Private, global	During reading, self-paced, scheduled
	Rating	External	Global	After reading, self-paced, scheduled
	Review	External	Global	After reading, self-paced, scheduled
	Discussion forum	External	Institution, global	After reading, self-paced, scheduled
	Tag	External	Global	After reading, self-paced, scheduled

List	External	Private, institution, global	After reading, self- paced, scheduled
------	----------	---------------------------------	--

Being part of the Amazon ecosystem, Goodreads affords a wide variety of DSR practices, starting with highlighting ebooks on the Kindle, which can then be displayed on one's own profile page, either privately or publicly visible, depending on the preferred settings. Beside highlights, all other kinds of user-generated content are external to the source text and in some cases allow to control who can see them, choosing between private, closed discussion groups, or global visibility. Tags are commonly called "shelves" on Goodreads, but on different pages of the platforms the terms "genre" or "tag" are alternatively used to refer to them. Lists are a more elaborate way of grouping together various books and creating meta-objects: they can include a long title and a description explaining the selection criteria; tags can be associated to the list; readers can vote books within a list; "like" it; and comment on it. "Challenges" or reading calendars organized by groups can make every aspect of content creation on Goodreads a scheduled activity, prompting the reading of books within a timeframe shared with others (weeks, months, a year) and, subsequently, influencing the pace of all other possible activities. In addition to social interaction, all users' actions on the platform are tracked and used to provide personalized recommendations, but they are also deliberately planned by users to organize and remember their reading history. That is why Goodreads is an example of DSR evaluation that supports both information management and purchase decision (DSR type 3).

3.4.1 Rating, reviewing, and their relation to readers' behavior and aesthetic judgement

People write book reviews for many different reasons and they choose on which platform to write according to their reviewing style and the kind of reviews they like to read or find useful. Regardless of the platform, "Customers generally contribute reviews for the purely cultural gains of altruism and recognition, and yet they also do so in full consciousness that their labor improves and extends a fundamentally capitalist machine" (Finn 2011, 69; cf. Steiner 2008). If not all of them, at least some users are certainly aware of the value, in terms of data, that they bring to these platforms, since each review contributes to improving the recommendation algorithm used to invite the purchase of more books.

Mike Thelwall, one of the scholars who dedicated more time to studying Goodreads, found that its users are around 76% women (Thelwall and Kousha 2017), and gender biases exist for both male and female readers (Thelwall 2019). Statistically reliable findings regard gendered preferences for literary genres and for books written by authors of the same gender as the reader (~500,000 ratings and reviews). For instance, on average, men rate short stories, horror, politics, and biography more highly, whereas women enjoy paranormal and contemporary romance more. But there are also differences within genres, like men rating male homosexual romance more highly than women, a fact that contrasts the preferences of AO3 readers. Both men and women like poetry written by female authors much more than that written by male authors, while the opposite is true for historical and contemporary romance, genres in which male authors are unanimously more appreciated. With respect to author-reviewer same gender bias, readers have opposite opinions on paranormal, erotica, manga, and classics. There are also cases in which only one gender shows a same-gender bias, like men who tend to like romance and YA fiction written by men much more than that written by female authors, or women who like adventure and historical fiction written by women. Thelwall's suggestion, based on the analysis of words frequency in reviews, is that attention to specific topics like relationships or war may be at least partly responsible for these biases. I think this is a plausible explanation but more research comparing the content of books and the gender/ratings of reviewers is needed to confirm this hypothesis.

More generally, over the years Thelwall's work has shown in a convincing way how online reviews can provide large-scale evidence about the reception of books (Kousha, Thelwall, and Abdoli 2016; Thelwall and Bourrier 2019; Bourrier and Thelwall 2020). Questions that need to be addressed to better understand readers who write reviews concern their social and educational status, as well as the intersection of these factors with preferences for certain authors or genres. Karen Bourrier and Mike Thelwall (2020) found that, for Victorian literature, the frequency with which books are taught in university courses predicts their number of ratings on Goodreads ($n = 203$; linear regression model, $b = .64$, $p < .001$, $R^2 = .47$). This suggests that many Goodreads users have probably attended a course in literature at a university and are thus well-educated, an insight consistent with data from other languages and nations (Van Putten-Brons and Boot 2017; Audunson and Tóth 2011).

If we look at large book clubs (50 Goodreads groups in English, with at least 500 members), rather than individual ratings, the situation is not very different in terms of variety: among the fifty books most read by book club members, almost half are works

of literature frequently taught at the secondary and postsecondary level (literary classics), and almost all are fiction (49), originally written in English (48), and being recently published and/or in the YA genre (Thelwall and Bourrier 2019). Moreover, “The 50 authors that are most popular amongst Goodreads book club members are mostly contemporary (28 alive in 2018), white (49) and from the USA or the UK (some originally, some after moving). A slight majority are male (28), although the top three are female and higher ranked authors tend to be female” (Thelwall and Bourrier 2019, 1145). Despite these commonalities, Thelwall and Bourrier acknowledge that typical reading is likely to vary according to book club theme and average age of readers. Focusing only on large book clubs leads to overlook an important affordance of digital platforms, namely the “long tail” of books and book clubs that are less popular in absolute terms but offer the opportunity for social interactions to readers who would be probably isolated otherwise (cf. section 1.3.3). Indeed, comparing their results with that of a 2018 survey of a representative sample of American readers ($n = 7,366$), Thelwall and Bourrier found that the list of authors most popular in Goodreads book clubs does not include authors who are not English or belong to ethnic, religious, or sexual minorities.

In light of this, questions to ask are whether digital platforms are standardizing the taste of readers by reinforcing conformism in aesthetic judgements and whether we can trust data coming from platforms with an underlying commercial interest, be it explicit or covert. We saw that educational institutions play a central role in determining the popularity of books, but it could also be that both retailers and publishers with great economic resources have a strong presence on digital platforms and succeed in influencing readers’ choices. An often-cited malpractice is that of buying fake reviews to promote or bury a book’s popularity and rating. However, although it is attested, it is difficult to assess how widespread it is and whether it has any important effect on readers’ behavior.

Nan Hu and colleagues (2012) studied the impact of allegedly manipulated reviews on books sales on Amazon.com (610,713 reviews of 4,490 books). First of all, they found that both the number of online ratings and their average score can reasonably explain most of the variability in the sales of the books: books that are reviewed more and rated higher are sold more (linear model, $p < .001$, $R^2 = .66$). Moreover, they also found a suspect pattern: the probability of seeing a positive review after a negative one was 2.6 times higher than the probability of a negative review following a negative review. Based on this, they concluded that around 10% of book reviews have been manipulated. However, they found no statistically significant effect of manipulated

reviews on book sales. According to other studies, which compared various kinds of reviews and platforms, a 10% estimate of fake reviews seems too high, a more accurate value is probably below 6% (Ott, Cardie, and Hancock 2012; cf. Mayzlin, Dover, and Chevalier 2014).

To better interpret these estimations, it is helpful to remind that, as a general tendency, people are more likely to write positive reviews than negative ones, people who had an ordinary or bad experience with a product often do not bother themselves to write about it. On the other hand, if our experience has been positive we are more likely to share it with others (Hu et al. 2017; Acerbi 2020; Hajibayova 2019). In the specific case of book ratings, Thelwall and Bourrier (2019) found that books that are most often read by Goodreads book club members are not necessarily liked or reviewed by them. The correlation between the average popularity of a book within book clubs and the total number of reviews is quite low (for the 50 most popular authors on Goodreads, $r_s = 0.40$, $p < .001$). Another related fact is that Amazon has more positive ratings than Goodreads (Dimitrov et al. 2015; Newell et al. 2016), but it does not necessarily mean that Amazon tweaks its rankings. The reason could be that their userbase is different from that of Goodreads, whose users do not all buy books online and are more critical of what they read (cf. Boot 2020). A similar trend holds true also for TripAdvisor (40% of 5-star reviews) and Expedia (50%) (Mayzlin, Dover, and Chevalier 2014). I recommend reading Acerbi (2020) for a better understanding of the behavioral dynamics of ratings and reviews, but, in summary: “The main challenge with online reviews is thus not that ‘trolls ruin everything’ or to recognize fraudulent positive or negative reviews, but to evaluate among too many upbeat reviews the ones that are only ‘lukewarm’” (Acerbi 2020, 93).

The relation between the verbal evaluation of a book and its rating could help in this task. This is still an understudied aspect, but Berenike Hermann, Thomas Messerli, and Simone Rebora presented some interesting results (Rebora et al. 2021). Analyzing the German-language platform Lovelybooks (1.3 million ratings and reviews), they found that words’ sentiment and users’ quantitative ratings are associated in a non-intuitive way. Sentiment is extremely positive only in five stars reviews, while reviews with one to three stars overuse negative sentiment. Four-star reviews are more nuanced,

expressing degree (somewhat, some, little), concession and limitation (however, nevertheless), as well as references to the ‘star rating system’ (deduction, four);

also, the diction more often refers to acts of criticism and deliberation (weakness, point of criticism), and, while friendly, appears more distanced (interestingly). (9)

Hermann, Messerli, and Rebora interpret this finding – and I agree with them – as evidence balancing the typical positivity bias for online reviews (Hu et al. 2017). A four-star rating may intuitively seem “positive” and a three-star rating “neutral”, but the diction used by DSR reviewers reveals that negative sentiment is overused for everything other than five-star ratings. Therefore, although ratings may seem suspicious at times, they should not be considered as mirroring the more complex reader response attested in reviews.

In a similar way, there is experimental evidence weakening the critique that online cultural consumption is making us more conformist or even more cognitively limited (Carr 2010). Psychological experiments have shown that the effect of popular choices and prestigious recommendation on aesthetic judgement is minimal. It is rather the perceived quality of cultural artifacts that mostly influences what people like. Acerbi (2020) discusses an experiment, among many others, in which participants listened to songs – while viewing either actual or manipulated charts ranking the songs and showing the number of downloads (Salganik and Watts 2008) – and notes that faking the ranking to promote certain songs led to a quarter fewer songs downloaded, because the “fake” top songs were downloaded less. The explanation is that the quality of the song (its rhythm and harmony) was still quite relevant for a decision directly related to cultural consumption and aesthetic experience, that is downloading the song for later use. Accordingly, since Amazon wants to sell us more books, it would be against their economic interest to tweak rankings and algorithms too much, unless they have a conspicuous instant revenue (e.g. in advertisement from publishers) that compensate for the huge loss in the long term (a potential 25% less profit from purchased books). I am pretty sure that Amazon does forecasting analyses that take into account all these factors.

I think we do not have to worry too much about being fed bad books or losing our cultural identity as readers because our taste is gradually conforming to that of everybody else. If that happens, I suspect it would be for other more complex reasons due to many interacting elements. On the contrary, the wide availability of easily retrievable niche products (the long tail) might be quite beneficial for many readers, since they can find books that satisfy their interests, ultimately bringing to an increase in cultural diversity, nurturing the passion for reading, and even offering chances to

develop skills related to critical reading and emotions while discussing with others. Looking at the content of reviews will allow us to cast more light on the latter aspect.

3.4.2 Reviews' content, emotions, and literary genres

From the perspective of professional literary critics, online book reviews have sometimes been considered an impoverished version of the reviews appearing in institutional journals and magazines, having very little value for the interpretation of literary works. Martin Rehfeldt (2017) opposed this view arguing that amateur reviews are better than professional reviews in showing the effect that a book has on its readers, since amateurs are not constrained by academic and editorial conventions that influence the selection of topics and the writing style. For instance, Ed Finn (2011) showed how reviewers on Amazon are hyperbolic in their evaluations, putting Forster Wallace's works in comparison with *Ulysses*, *Moby Dick*, and even *Hamlet*, rather than with works of other great writers of his generation, like professional critics prefer to do. In other words, for Amazon reviewers, Wallace is ready to enter the canon of English literature faster than for professional critics.

Various researchers have worked on identifying the topics of amateurs' online book reviews, using different methodologies (Mehling et al. 2018; Hegel 2018; Hajibayova 2019; Rebora et al. 2019; Boot and Koolen 2020; Koolen, Boot, and van Zundert 2020; Walsh and Antoniak 2021). Unfortunately, the results are difficult to compare because each study starts from different theoretical premises, different definitions of concepts, use different methods, and some relies on manual annotation of reviews, introducing a further element of arbitrariness. Many of these works show how controversial it is for humans to clearly identify the expression of a mental or affective state in a sentence, and how the lack of accuracy of algorithms is often due to a semantic indeterminacy revealed by humans in the first place. However, regardless of the different frameworks, it may still be useful to look at these results together, so I grouped them in Table 11. Rather than comparing values from different studies, it is more sensible to look at proportional relations between the themes identified by each study, and see whether they vary across other studies and platforms. That is, compare the numbers in the table vertically, before comparing them horizontally. To facilitate this, I calculated the z-scores of the occurrences identified by each study, i.e. the variation of each value with respect to the mean occurrence of all themes identified by that study. Thus, negative values indicate themes that occur less frequently than average and positive values indicate themes that occur more often than the mean frequency.

Table 11. Percentage of Dutch, English, and German online book reviews talking about a specific type of content

DSR platform	Amazon.de: German (Mehling et al. 2018)	Goodreads: English (Hajibayeva 2019)	Goodreads: English (Hegel 2018)	Goodreads: English (Rebora et al. 2020)	Hebban, Bol, and others: Dutch (Boot and Koolen 2020)	Hebban, Bol, Goodreads, and others: Dutch (Koolen et al. 2020)	Professional reviews: English (Hegel 2018)
Summary	30% (+0.24)	25% (-0.55)	-	-	-	-	-
Plot	32% (+0.33)	-	22% (+1.74)	-	-	-	37% (+1.81)
Characters	21% (-0.16)	7% (-1.41)	12% (+0.6)	-	-	-	20% (+0.50)
Narrative style	23% (+0.7)	52% (+0.73)	-	-	-	-	-
Style/language	8% (-0.75)	-	5% (-1.13)	-	-	-	6% (-0.58)
Overall appreciation/dislike	79% (+2.45)	59% (+1.07)	13% (+0.23)	-	35% (-1.09) ¹	36% (-0.90) ¹	8% (-0.42)
Identification/empathy	7% (-0.80)	40% (+0.16)	11% (-0.11)	1.3% (-0.71)	65% (+0.87)	40% (-0.18)	8% (-0.42)
Suspense	46% (+0.96)	-	-	0.8% (+0.71)	-	-	-
Emotions	-	-	7% (-0.79)	-	55% (+0.22)	47% (+1.08)	2% (-0.89)

1. Includes appreciation only, not dislike.

Data are taken from six different studies investigating online reviews' content. In some cases, the labels reported for the type of content are different from those originally used by the authors. Boot and Koolen's category "narrative feeling" includes both identification with the characters and response to other narrative effects like suspense. I did not include all the categories identified by each study, just those used in more than one study. Within brackets are the z-scores, i.e. the variation of each value with respect to the mean occurrence of all themes identified by that study.

In [Table 11](#) there is also a column for professional reviews taken from several online newspapers and magazines (Hegel 2018), which can be used as a reference to compare the different attention given by amateurs and professionals to the same topics. In the case of Allison Hegel's work, we can compare these values directly, since they have been obtained using the same methodology. When reviewing books, English literary critics use a substantial part of their words to describe the plot and the characters, talking very little about the emotional impact of the text. On the contrary, amateurs spend 15% less words for the plot and 8% less for characters' description, but 3-5% more to express their evaluation of a book, the identification with characters, and other emotions elicited by a story. On the German version of Amazon, only average attention is given to plot (z-score = 0.33); Gabriele Mehling and colleagues (2018) link this to the fact that the initial reviews of a book usually summarize the content, so later reviewers do not bother to do that again, since it would be a repetition for other readers (indeed, this is explicitly mentioned in some reviews).

Evaluating the quality of a book in terms of overall appreciation or dislike is a central aspect for both English and German amateur reviewers. However, it seems that Dutch readers are less concerned with expressing their positive evaluation of a book

(“aesthetic feeling”) and rather prefer to share their emotional response to the story and to the characters (“narrative feeling” and “emotional impact”) (Boot and Koolen 2020; Koolen, Boot, and van Zundert 2020). More generally, the style or the quality of an author’s writing are not mentioned often by reviewers, not even professionals, who probably deal with them in more detail when writing critical articles rather than reviews. There seem to be also other peculiarities dependent on language and culture: for instance, English and Chinese tend to use more assertive expressions to push other readers to carry out their recommendation (e.g., “Try him!”), whereas Japanese reviewers tend to write about their personal feelings, instead of directly and explicitly recommending a book (Chik and Taboada 2020).

Marijn Koolen and colleagues also looked at the co-occurrence of topics within book reviews and found that narrative feelings correlate with aesthetic feelings, an association which is higher for readers who are more prolific (for users who wrote more than ten reviews, $r = .32$, $p < .001$). Very active reviewers are more balanced in writing about both narrative effects and their appreciation of the book, whereas occasional reviewers are more likely to mention only one of the two topics. Interesting insight can be obtained from the computational analysis of reviews, but it is always tricky to generalize the results, because many variables affect both the automated detection of topics and the reviewer’s expressive preferences. For instance, it is easier to extract topics from longer reviews, frequent reviewers are more balanced in writing about different types of impact, and readers of some genres seem to have their peculiarities (Koolen, Boot, and van Zundert 2020).

So far, I have tried to elaborate on the variability of the content of amateur reviews, but it is now time to talk about the elephant in the room: literary genres. Genre is the concept around which the majority of DSR practices are organized, and reviews on platforms like Goodreads show how distinctions between genres shape readers’ discussions.

Goodreads creates and features genre-based book clubs, discussion groups, and “Favorite Genre” sections in every user profile, assuming that its users want to primarily identify and interact with other genre fans rather than forming groups and friendships based on different characteristics, such as their nationality or age. (Hegel 2018, 100)

The influence of genre on reviews has been considered by many of the researchers I mentioned. Dutch readers of science fiction tend to write longer reviews than readers of other genres, and reviewers of teen fiction use much more first and second person

singular pronouns than average, e.g. “it will leave *you* speechless.” The latter phenomenon could be due both to the writing style – assuming that readers of this genre are younger, which is not to be taken for granted – and to the fact that this genre usually emphasizes emotions (Cart 2016), which are then discussed in reviews. Another specificity due to genre regards the far greater occurrence of narrative feelings (and the concurrent limited appearance of aesthetic feelings) for thrillers, whereas the opposite is true for generic fiction (Boot and Koolen 2020). Dutch readers of thrillers are less likely to talk about stylistic features, focusing more on the core dynamic of the genre – suspense generation – which, if present, is probably a guarantee of appreciation.

With respect to German readers, we know that their reviews of romance are characterized by an overuse of terms related to joy and a limited use of fear-related terms, in comparison to the average of all genres; while the opposite is true for reviews of thrillers, with the addition of an abundant use of surprise-related terms. Interestingly, in reviews of general fiction, terms related to surprise are used much less than average and there is a wide use of words related to sadness (Messerli, Rebora, and Hermann 2020).

For English reviews, an extensive work has been done by Hegel (2018), who not only showed that reviews of different genres have different dominant topics, but also that different communities and cultures of readers are concerned by the discussion of genre in various degrees. By looking at the words appearing directly beside genre-related terms (collocations) for both professional and amateur reviews, Hegel concluded that for professional critics genre is something intrinsic to the book and only a few genres are considered when deciding which book to review. On the other hand, for amateurs genre is something defined by a way of relating to the story, and they are much more flexible, and detailed at the same time, in diversifying various genres.

Professional reviewers, when they discuss genre in terms of its structure, treat genre as a characteristic of books. Amateur reviewers, however, with their more personal view of genre as a way to shape identities and form communities, reveal a shift in genre that’s occurred since reviewing has opened up to the public. [...] amateurs show this new way of defining genre, as a personal category of fandom that is then negotiated to determine whether or not it applies to a given book. By comparing professional and amateur book reviews, we can see genre’s transition from literary category to consumer category firsthand. (Hegel 2018, 40)

Based on how Goodreads users use tags/shelves to classify a book as belonging to a given genre, Hegel observes an “emergent landscape of genre based on readers’ actual use of genre labels rather than pre-existing or theoretical genre hierarchies” (80). Users can freely create new tags and assign them to books, creating a new piece of information that will be displayed on the book’s webpage, if the tag becomes popular enough, also showing how many users assigned it to the book. Coherently with what observed in the text of reviews about the relational conception of genre for amateur readers, Hegel looks at how genre differentiation emerges through the use of reader-assigned tags. The procedure consists in grouping tags together based on the number of books they have in common (18,239 tags; pipeline: cosine distance, t-SNE dimensionality reduction, hierarchical clustering). In other words, two tags/shelves/genres are similar if many users assigned both of them to the same book, while they are more different from each other if no user used them for the same book. A clustering algorithm calculates the distance between tags based on their similarity, so it is possible to plot them on a bidimensional plane and look at what tags appear closer to each other, possibly forming groups that can be interpreted as user-defined genres.

In some cases, many readers assign the same pairs of tags to the same books pretty often and consistently, so groups of tags very close to each other can clearly be identified as a genre, like in the case of “drama,” “theater,” and “plays.” But in the majority of the cases the boundaries between genres are less clear or they challenge established conceptualizations. An example revealing the fluidity of genre for lay readers is that of “horror” (Figure 11). Hegel marked in orange the tags that, according to her, are more representative of the horror genre, but she also notes that “angels” and “shorts” are much closer to “horror” than “terror” is, suggesting that “horror as it functions on Goodreads isn’t as much about inspiring fear as it is about tropes like ghosts and angels as well as an abbreviated format” (112). Thus, a method like clustering can reveal characteristics of a genre but, since tags are a tool used for information management by readers who want to organize their virtual libraries, we can also interpret these results in a more pragmatic way. In this example, we can notice that for readers it is important to remember that a horror story is short and features ghosts or angels, but also that they probably rely on information institutionalized by Goodreads, since “Horror” is an official genre for which a description is provided, while the page for “Terror” only lists books that have been tagged as such by the users.¹



Figure 11: Close up of the cluster for the horror genre, identified using a machine learning algorithm calculating the similarity between tags based on which books they have been assigned to.

See Hegel (2018, 109) for the whole figure including the other identified clusters.

Besides using tags, the importance and cohesion of genres for readers can be also explored looking at the text of reviews. Overall, Goodreads reviewers discuss genre significantly more ($p < .05$) in science fiction reviews (~5.8% of total reviews) than in romance (~2.3%), horror (~4.1%), and mystery (~4.4%), showing that genre and subgenres in some cases are more important than in others. Moreover, in the case of sci-fi, readers are concerned about the status of the genre and defend it against its detractors; for this reason, they “consistently use the language of awards and value to lend science fiction literary legitimacy” (Hegel 2018, 57). This kind of evidence shows the apparent unavoidability of the power relations of social distinction described by Pierre Bourdieu (1984) and John Guillory (1993): the internet is a democratized space only to a limited extent, amateur book culture still seems to be ruled by the influence of institutionalized authorities like awards committees (cf. Long 2003; Finn 2011; Rebora et al. 2021). Not necessarily because their opinion is the right one, but because it is an opinion that matters and as such can be questioned (see some examples in Hegel 2018). Opposing and alternative reading tastes and conceptions of genre exist only in relation to the dominant literary culture (R. Williams 1973). An additional proof of this is that none of the 70 popular science fiction books sampled by Hegel (2018) was labeled by more than 1% of the readers with tags related to both awards and emotions. In other words, lay readers enact the institutionalized incompatibility between books that are worth remembering because they won awards and those worth remembering because they elicited emotions (cf. Piper and So 2015).

Returning to the analysis of the reviews' text, Hegel selected a random sample of 262,900 Goodreads reviews in English, posted in the time range 2007-2016, keeping only those explicitly tagged as belonging to a genre. Starting from this dataset of labeled reviews, it is possible to identify the most distinctive words used in reviews of each genre (TF-IDF). In turn, such groups of genre-specific terms can be used to check whether there is any change in the frequency with which they appear in reviews over time (cf. K. Chang et al. 2020; and Sharma et al. 2020, who did something similar with professional reviews over a longer time span). This method allows to check whether the reader response verbalized in reviews changes over time, namely whether it becomes more consistent or more varied. To do so, we can teach a predictive machine learning algorithm to recognize reviews as belonging to a certain genre based on the presence of genre-specific terms: the more a review contains words specific for a certain genre, the more likely it is that such review is discussing that genre. But we are not really interested in the accuracy of the prediction itself, we rather want to observe whether it changes if we target reviews posted in different years, because the accuracy of the prediction is based on the presence of certain recurring genre-specific terms. Hegel found that Goodreads reviewers grew noticeably more predictable in how they talk about genre over time, i.e. they used an increasingly more similar vocabulary within reviews for each genre. This result may look like a victory for the detractors of digital reading, amateur reviews, and internet conformism in general (cf. section 5.2.1), but if we look at the accuracy values of predictions for both amateur and professional reviews, we see that literary critics have been far more predictable since 2007. They are identifiable by the algorithm with an accuracy always higher than 85% (F1 score), whereas the predictability of amateur reviews increased from 64% to 73% in ten years. This suggests that professionals strictly adhere to a common jargon – and probably also text structure – when writing reviews.

In the 21st century, professional and academic reviews are a consolidated practice, therefore it is not surprising that they use a more cohesive language. They are trained in universities where established conventions are transmitted, and these get enacted when writing. Applying cultural evolution theory (Mesoudi 2011; Morin 2015) to amateur reviews can explain what happens to a discursive practice over time: thanks to digital media, amateur reviews spread widely (more availability) and useful reviews started to be copied more often (higher fidelity of transmission of discourse-specific traits) (Acerbi 2020), also thanks to Goodreads algorithm, which displays on the first page the reviews that have been marked as most useful by users. As a consequence of this high availability and fidelity of transmission, the language of reviews became

overall more similar. This is not only a general phenomenon, but something observable within genres and subgenres as well, whose reviews became in time more and more specific in discussing traits perceived by readers as the most relevant for (typical of?) each genre (Hegel 2018). Different communities are characterized by the transmission and evolution of different practices: for instance, among literary scholars, emotional language did not get selected as one of the traits to be transmitted; among Goodreads amateur reviewers, it did. Actually, we have evidence that for literary scholars this bias is not only related to book reviews, an excess of emotions and sentimentality in fiction itself is deplorable, too, according to them (Piper and So 2015; Borghesi 2018).

If we look at the most frequent words used nearby the mention of genre, we can see patterns that show how amateurs talk about books in different ways. I already mentioned the importance of the status of the sci-fi genre for its readers; another notable difference between amateurs and professionals regards romance (Hegel 2018, 63-67). The former focus more on further specifying details of a book in relation to genre – e.g. with words about settings, themes, or subgenres (“paranormal,” “historical,” “college,” “cheesy,” and “erotic”) – whereas the latter often provide contextual information, e.g. mentioning words like “debut” and “adult,” providing a service helping potential readers to be aware of the book’s positioning within the publishing landscape. Moreover, if we look separately at negative (1-2 stars) and positive (4-5 stars) reviews, we can note that family-related terms occur significantly more frequently in professionals’ negative reviews of romance, compared to positive reviews or amateur reviews. Professionals also consider “romantic,” “plot,” focus on “self,” and being for “women” to be negative features (cf. Felski 2009; Radway 1984).

It is possible that the way in which reviewers talk about books is influenced by the tacit authority of Goodreads’s infrastructure – being the apparatus enabling the existence of a universe of amateur reviewers. Hegel showed that the language of reviews became more predictable over time for all genres: a possible explanation is that Goodreads is leading reviewers to use certain words and focus on certain topics in their reviews, by explicitly listing subgenres on the main webpage, by codifying genre tropes as cataloguing categories (like “aliens” and “robots” as science fiction subgenres), and by displaying the most “standardized” reviews at the top of the list for a book. All these platform features contribute to consolidate a certain way of talking as exemplary for a genre. This is confirmed by the fact that the predictability of the reviews showed after the first webpage of results decreases drastically (Hegel 2018, 139-141).

On one hand, the fact that the homogeneity of reviews is based on tropes is useful for readers looking for books recommendations, since they can more easily understand the plot, characters, and themes of a book, and decide whether to read it or not. On the other hand, Hegel's interpretation is that this means a progressive reduction of discourse diversity. However, more intra-genre cohesiveness does not mean that readers will not find information about evaluation or sentiment, since these topics are present in many reviews (they are not distinctive in terms of TF-IDF). Rather than seeing Goodreads as encouraging "tight-knit communities of people who all generally agree and are similar ... less likely to be exposed to surprising ideas or to broaden the range of possible responses they associate with the genre" (Hegel 2018, 131), I think it evolved into a very efficient and informative system, allowing to discuss evaluation and emotional response together with genre-specific tropes.

Rita Felski's words can be useful to summarize the specificity of amateur reviews:

Scholars like to think that their claims stand or fall on the merits of their reasoning and the irresistible weight of their evidence, yet they also adopt a low-key affective tone that can bolster or drastically diminish their allure. Critical detachment, in this light, is not an absence of mood, but one manifestation of it—a certain orientation toward one's subject, a way of making one's argument matter. It is tied to the cultivation of an intellectual persona that is highly prized in literary studies and beyond: suspicious, knowing, self-conscious, hardheaded, tirelessly vigilant. (Felski 2015, 6)

While some amateur readers may emulate this style, Hegel's work shows how they are more free to express the "sensuous and somatic" side of their aesthetic reaction (Felski 2009, 76). Needless to say that literary scholars feel passions as well while reading, although they are reluctant to publicly talk about it.

As a final note, before moving to the next chapter, let me point out that, besides thematic differences, reviews by amateurs and professionals can also be told apart because of their style. Comparing Italian reviews published on aNobii (a platform similar to Goodreads), a newspaper, and academic journals, Rebora found that the different number of occurrences of the 2,000 most frequent words – thus including many functional words like pronouns and prepositions, which are normally excluded in the analysis of themes – is a feature more effective than content when predicting via machine learning to which category a review belongs (Rebora et al. 2021). More in-depth research is needed to better observe the nuances of such stylistic differences

and identify the characteristics of amateur reviews, which will probably vary with the genre discussed, and maybe also with the nationality and gender of the reviewer.

Footnotes

1. <https://www.goodreads.com/genres/horror> and <https://www.goodreads.com/genres/terror> ↵