



The role of processing foregrounding in empathic reactions in literary reading

Giulia Scapin, Cristina Loi, Frank Hakemulder, Katalin Bálint & Elly Konijn

To cite this article: Giulia Scapin, Cristina Loi, Frank Hakemulder, Katalin Bálint & Elly Konijn (2023) The role of processing foregrounding in empathic reactions in literary reading, *Discourse Processes*, 60:4-5, 273-293, DOI: [10.1080/0163853X.2023.2198813](https://doi.org/10.1080/0163853X.2023.2198813)

To link to this article: <https://doi.org/10.1080/0163853X.2023.2198813>



© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.



[View supplementary material](#)



Published online: 01 May 2023.



[Submit your article to this journal](#)



Article views: 1838








[View related articles](#)



[View Crossmark data](#)

The role of processing foregrounding in empathic reactions in literary reading

Giulia Scapin ^{a,b,c}, Cristina Loi ^{c,d,e}, Frank Hakemulder ^{c,d}, Katalin Bálint ^a, and Elly Konijn ^a

^aDepartment of Communication Science, VU Amsterdam; ^bDepartment of Communication, Haifa University; ^cInstitute for Cultural Inquiry, Utrecht University; ^dNational Reading Centre, University of Stavanger; ^eDigital Humanities Lab, University of Basel

ABSTRACT


A considerable body of research has examined the age-old assertion that reading literature enhances empathy, however, mixed results have been found. The present study attempts to clarify such disparities, investigating the role of foregrounding in possible differences in readers' processing of literary texts and its connection with readers' empathic reactions. We asked participants ($N = 78$) to mark parts of the text they considered as "foregrounding" (i.e., deviating from "normal" discourse), and we analyzed how they processed these stylistic aspects. Participants' open responses to one of two selected texts were categorized as either Shallow, Failed, Partial, or Full Processing of Foregrounding. Full processing was associated with higher Comprehensive State Empathy Scale scores than Failed Processing. Stylistic analysis of word combinations that participants marked as "striking" suggests that, rather than stylistic devices per se, readers' depth of processing may enhance state empathy.

Literature is accredited with a range of powerful functions. One prominent idea is that it may enhance readers' ability to empathize with people different from them (Hakemulder, 2000, 2020; Koopman & Hakemulder, 2015). Potentially, this effect could have a valuable impact, for instance, in reducing interpersonal conflicts and societal tensions associated with prejudice and stigmatization (cf. Johnson, 2013). It is perhaps because of such promises that the past decades have seen a growing number of studies testing these assumptions (Calarco et al., 2017; Dodell-Feder & Tamir, 2018; Quinlan et al., 2022). Unfortunately, mixed results have been found (see review in Wimmer et al., 2021). Moreover, meticulously executed attempts to replicate positive effects have failed (e.g., Samur et al., 2018). Because such disparities are hard to explain, some researchers have called for a change in the research agenda: the field should move away from the question of whether or not reading literary texts enhances empathy but rather aim to establish when it does and when it does not (De Mulder et al., 2022; Mumper & Gerrig, 2017). The present paper aims to contribute to this objective by using an operationalization that differs fundamentally from previous research.

What is literature?

The question of "what *is* literature?" is a notoriously contentious issue, both inside and outside academia. Previous researchers frequently ignored the matter altogether, for example, by focusing

CONTACT Giulia Scapin  g.scapin@vu.nl  Department of Communication Science, VU Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/0163853X.2023.2198813>

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

on exposure to narrative fiction rather than on literature in particular (e.g., Mar et al., 2009). However, not all narrative fiction is considered literature (e.g., Dan Brown's 2003 novel *The Da Vinci Code*), and not all literature can be categorized as narrative fiction (e.g., most poetry is not, nor are literary autobiographies). Others refer to an assumed consensus about what it may be (e.g., by using texts written by acclaimed authors lauded with prestigious prizes; Kidd & Castano, 2013). But whether an author is renowned and respected as a literary artist does not allow us to identify what precisely in the processing of their texts might be responsible for enhancing empathy. Also, literary fame can be volatile and susceptible to factors that do not necessarily guarantee literary quality. According to some literary critics and sociologists, it is mainly ideological considerations of "gatekeepers" like editors and publishers, as well as authors' gender and ethnic background that determine whether a text is included in the literary canon (Berkers, 2009; Herrnstein Smith, 1988; Janssen, 2001; Verboord, 2003).

So, what should be our definition of literature? To avoid discussing in detail the debate about what makes a text literary, we can best approach it as follows: one side of the debate argues that literary quality can be identified in text qualities that differ from "normal" discourse; some examples of these text qualities are unusual metaphors, infrequent words, ungrammatical sentences, parallelisms and all other rhetorical devices that act on the level of discourse and semantics or what we will refer to as foregrounding, textual foregrounding, foregrounded devices, stylistic deviation or distortion (Hakemulder, 2020; Simpson, 2014; Van Peer et al., 2021). The other side of the debate uses sociological and constructivist arguments, suggesting that what is classified as literature depends on historically unstable conventions determined by power relations in a society and that it is the cultural capital of a particular elite group in society that allows them to legitimize literary status (e.g., Ahearne & Speller, 2012). In the present context, we do not need to take a stance in this debate. Still, it seems clear that the first definition of literature is most relevant for the current research: the theoretical claims that reading literary texts enhances empathy is related directly to its unique stylistic features and the processing of these features (Koopman & Hakemulder, 2015; De Mulder et al., 2022), and not to how a taste for specific texts correlates with the cultural capital of a certain status group.

The studies we discussed until now did not directly test the role of text factors in affecting empathy. In fact, some researchers did take textual definitions of literature as a starting point, focusing on those text qualities that they considered to be distortions of normal language usage (e.g., Kuzmičová et al., 2017). This kind of research, then, examined whether those aspects were responsible for readers' empathic responses. However, this approach ignored whether readers were actually aware of the norms that were distorted or whether they noticed the distortions. Consequently, it remains uncharted how readers act upon that awareness. It seems plausible that such an awareness is preconditional for literary effects to occur. Whether a deviation (e.g., a grammar rule or genre convention) becomes prominent in readers' processing of the text may depend on whether readers have access to what is the norm in the first place (cf. Hakemulder, 2020). Here, we would like to argue for an approach that centers on readers' processing of the texts, that is, what they themselves notice as being unusual in the text, how they respond to those elements in the text, and how the various modes of responses to foregrounding are related to empathy.

Empathy and literature

Literary reading is assumed to enhance empathic reactions in the reader by simulating social cognitive activities similar to those in real-life social situations (Oatley, 2012). According to Mar (2018, p. 454), "stories could bolster social cognition either through (1) frequent engagement of social-cognitive processes or (2) the presentation of explicit content about social relations and the social world." Because not all stories are literature, the question is, "why literary reading would have a special role in eliciting empathic reactions?" One option to clarify the impact of literature on empathy is to take a closer look at what literary scholars propose the effective ingredient might be. The text quality we referred to above as foregrounding seems a likely candidate for disentangling the relationship between literature reading and empathy (e.g., Miall & Kuiken, 1999). As proposed by the *dehabituation model*

(Fialho, 2019; Miall, 2006), distortions in the text (i.e., foregrounding devices) obstruct automatic text processing, prompting readers to reflect and become actively engaged in finding a coherent understanding of the text, which then leads to deeper insights. These insights supposedly induce a transformative effect, such as a (re)appraisal of cognitive schemata (cf. refamiliarization in Shklovsky, 1917/2016, p. 80; cf. schema refreshment in Cook, 1994; Hakemulder, 2020). Readers would reconsider their perspective and feelings, as described by the linguist and literary theorist Shklovsky (1917/2016, p. 80): “to make the stone stony.”

This reconsideration does not apply only to objects (i.e., the stone) but also to people presented in the story world, helping the readers to consider the point of view of the characters (i.e., perspective-taking) and feel with and for them (i.e., empathic reactions), upsetting “the stereotypical schemata” we use to navigate the world (Koopman, 2016, p. 64). For example, when we read this sentence from *The Bell Jar* by Sylvia Plath (1963/2013, p. 3) “I felt very still and very empty, the way the eye of a tornado must feel, moving dully along in the middle of the surrounding hullabaloo,” we are discovering a new perspective, a new sensation connected with the concepts of *stillness* and *emptiness*, and it helps us to understand better the inner world of the story character presented.

Controversial results

Literary reading is usually associated with a longer reading time (e.g., Miall & Kuiken, 1994; Zwaan, 1991), supposedly allowing readers to reflect on a deeper, non-literal meaning (cf. Hakemulder, 2020). Experiments by Kidd and Castano (2013) did suggest a specific positive effect of *literary* fiction, in contrast with popular fiction, nonfiction, and not reading, on understanding others’ mental states. However, as mentioned earlier, replications of this effect were not always successful (e.g., Samur et al., 2018), whereas others were (e.g., Van Kuijk et al., 2018). Furthermore, these studies observed groups of participants assigned to different conditions but without considering the potential diversity of processing of the literary texts among readers.

Other attempts sought to connect foregrounding to enhanced empathic reactions more specifically but found contrasting results. For example, in the study of Koopman (2016), participants were exposed to three different versions of a literary text: *original* (e.g., “The farewell. Carrying the body to the burial. Seeing it off. Carrying. Setting up the place where she would be from now on. Taking possession of the cemetery as an outside living room”); *without imagery* (e.g., “The farewell. Carrying the body to the burial. Seeing it off. Carrying. Setting up the place where she would be from now. Staying at the cemetery constantly”); and *without foregrounding* (e.g., “And then the farewell, with the carrying of the body and seeing it off. They took her daughter to the place where she would be buried. She would go to that cemetery very frequently”; Koopman, 2016, p. 86). Results supported a positive effect of foregrounding on empathic responses, defined in Koopman’s study as feeling and understanding the character’s feelings and situation.

Another study by Kuzmičová et al. (2017) suggested an opposite direction of the relationship between foregrounding and empathy. The authors applied a qualitative approach, exposing participants to two versions of the same text: the *original* (e.g., “As a matter of fact he was proud of his room; he liked to have it admired, especially by old Woodfield”); and the *nonliterary* version, where foregrounding was reduced (e.g., “He was proud of his office; he liked when people admired it, especially old Woodfield”; Kuzmičová et al., 2017, p. 142). The nonliterary version elicited more empathic responses than the original, literary version. It may be that some stylistic devices are more consequential than others: some may stimulate readers to reconsider their judgments about characters and thus generate more empathy; others may cause readers to distance themselves from the text and the characters and thus hinder empathy.

The design and approaches applied in the previous studies do not allow any specific conclusions about stylistic devices’ functions. Therefore, in the present study, we will explore the relations between various forms of textual foregrounding and the readers’ experience of empathy. A second potential factor is a variety in readers’ responses to foregrounding. Because foregrounding depends on the

background (i.e., readers' access to knowledge about the conventions and grammar rules, etc., that are distorted; cf. Hakemulder, 2020), and because that background may vary from reader to reader, it seems essential to consider the individual experience of the reader. Moreover, not all readers may be equally interested in strategizing to interpret the distortions (e.g., Koek, 2022), with some preferring a more cursory processing of such potentially annoying obstacles, while others are curious to find a non-literal meaning. Consequently, we are still left in the dark as to how and when reading literature is associated with enhanced empathic reactions. We suggest this is because we do not know whether the *literary* aspects of the texts were processed at all, nor whether such processing could be associated with enhanced empathy.

Considering failed foregrounding

The research we have explored so far starts from the assumption that readers process the text similarly. They go through a path, as proposed by Harash (2019, 2021), divided into three main stages to successfully process stylistic distortions present in a literary text. These stages are the following: 1) the distortions in the text hinder the automatic processing of the message included in the text (*defamiliarization*, e.g., unclear meaning, lacking info, unexpected elements; including distortions at the semantic level, events that do not tally with readers' real-world expectations); 2) readers are then stimulated to reflect and make an effort, playing an active role in the interpretation; 3) this effort results in the appreciation of the artistic abilities of the author and in the reappraisal of the personal schemata related to the contents presented (*refamiliarization*; cf. dehabituating model, Miall, 2006; Figure 1). For example, reconsider the sentence from Sylvia Plath discussed above ("I felt very still and very empty, the way the eye of a tornado must feel, moving dully along in the middle of the surrounding hullabaloo," 1963/2013, p. 3). Readers may recognize a combination of images that are not frequently used in everyday language; still, these images may be interesting and intriguing, stimulating the reader to think about a deeper reason for the choice of these words, which convey the character's struggle.

In his theorization, Harash proposed a model of Failed Foregrounding (Harash, 2019, 2021) where the abovementioned process is not always successful; in fact, the reader might experience an incomplete process and stop at one of the stages described before. Through his analyses of readers' thinking-aloud responses and eye-tracking data, Harash defined four different types of elaboration of the stylistic distortions: *shallow processing*, where distortions are not processed at all (i.e., the readers stop after stage 1); *failed processing*, where the reader tries to make sense of the distortions but without arriving at a solution, triggering frustration and negative aesthetic experience (i.e., the readers stop after stage 2); *partial processing*, where

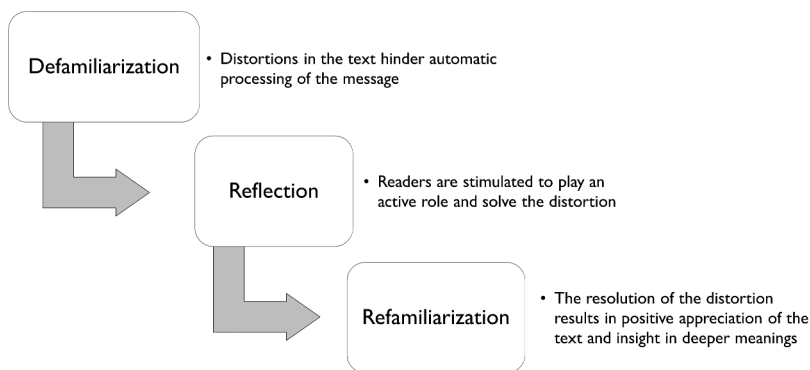


Figure 1. Schematic representation of the three main stages to successfully process stylistic distortions in a literary text (adapted from Harash, 2021).

the reader attempts to find meaning in the distortions but without arriving at a full conclusion, and unresolved discrepancies are associated with a positive aesthetic experience (i.e., the readers stop after stage 2 as well); and *full processing*, where the reader has an insight regarding a deeper meaning related to the distortions, and which is associated with a positive aesthetic experience (i.e., the readers complete the whole path successfully).

The present study

The present study explores the relationship between stylistic distortions that readers notice (also called “perceived foregrounding” in previous literature; Hakemulder, 2020; Van Peer et al., 2021) and empathic reactions by applying a different approach from the previous research. Data collection and analysis focused on the depth of readers’ processing of those distortions that they detect themselves, in contrast to earlier studies in which the researchers determined the presence of such distortions (cf. the concepts of perceived foregrounding vs. textual foregrounding in Van Peer et al., 2021). To do this, we used Harash’s (2021) categorization of the various ways in which readers process stylistic distortions. The central hypothesis of this study is that readers’ deeper (full) processing of stylistic distortions is associated with higher empathic reactions toward the story character. To our knowledge, this is the first attempt to operationalize readers’ processing of stylistic distortions and their link with empathic responses. In addition to the main variables, the study also included an assessment of text appreciation and reading expertise. The aim of including such variables was to check the consistency of the present design with the design of Harash’s study (2021). In fact, in previous findings, Full Processing was connected with higher appreciation and higher reading expertise.

The present explorative study builds on another component introduced in Harash’s dissertation (Harash, 2019), namely the possibility of using the Failed Foregrounding Model to understand the effectiveness of stylistic devices in completing the foregrounding process. Thus, a stylistic analysis has also been performed, investigating the textual features that readers have identified as striking coherently with a bottom-up approach, that is, starting from what the readers perceived as foregrounding. We examined the emergence of possible patterns in the groups of Processing of Foregrounding, categorized according to their level of processing of stylistic distortions. Such exploration might help understand the relation (or the lack thereof) between foregrounding and empathy.

The present study is part of a larger project (Scapin, 2022), which focuses on the potential of literary reading to enhance an empathic reaction toward a character who presents depressive behavior. The main aim is to understand the potential of literary reading in eliciting empathy toward people living with depression and reducing the stigmatized attitudes related to them.

Method

Participants

Seventy-eight participants were recruited, including 47 females, 30 males, and 1 non-binary (for statistical reasons, this participant was randomly reassigned to one of the other two groups; Bálint et al., 2022), with an age range from 18 to 76 years old ($M = 38.73$, $SD = 14.05$). Participants were pre-selected through an online platform (Prolific) using the following inclusion criteria: native speakers of English, a high school diploma or higher, and an interest in literature. These criteria were considered necessary to comprehend the textual stimuli well. Data were collected through an online Qualtrics survey. Informed consent was collected from all participants, and a debriefing was provided at the end of the survey. This study was approved by the Research Ethics Review Committee of the Graduate School of Social Sciences of the VU Amsterdam. The present study was not pre-registered.

Procedure

Participants were randomly assigned to read one of the two texts, *The Bell Jar* by Sylvia Plath ($n = 38$) or *Stars and Saints* by Lucia Berlin ($n = 40$). After the first reading, participants were asked to complete the Comprehensive State Empathy Scale. Then, they were asked to read the text again and highlight the parts they perceived as “unexpected, unfamiliar, different or disruptive”; afterward, they responded to the open question about their experience. Finally, they filled out the appreciation and reading expertise questions (see Measures section below).

Materials

Two texts were used as stimuli: an extract from the novel *The Bell Jar* by Sylvia Plath (1963/2013, 592 words) and one from the short story *Stars and Saints* by Lucia Berlin (2016; 954 words, full texts are available in Supplementary Materials). Initially, a total of 16 texts were evaluated by three experts in literary studies, who were asked to rate the pertinence of the topic from 1 to 7 (i.e., to which degree they thought the texts talk about depression), and whether the texts should be considered examples of good literature (from Dixon et al., 1993). Of those 16 texts, Plath’s and Berlin’s extracts received high and comparable rankings on the pertinence of the topic (Plath’s had a mean of 5 and Berlin’s of 6); and both texts were rated similarly as examples of good literature (Plath’s had a mean of 4.75 and Berlin’s of 4.33). No significant differences were found between the texts on both criteria (for the topic: $t(4.99) = 0.31$, $p = .77$; for the rating as good literature: $t(4.74) = 0.93$, $p = .40$). Another group of experts in foregrounding and stylistics ($N = 7$) rated the overall level of foregrounding (“On a scale from 1 (very low) to 10 (very high), how much do you evaluate the overall level of foregrounding of this text?”) presenting a significant difference (Wilcoxon rank sum test $W = 46$, $p = .006$) where the extract from Sylvia Plath had a higher score ($M = 7.43$, $SD = 1.27$) than the extract from Lucia Berlin ($M = 4.28$, $SD = 1.5$). However, no agreement was found on which parts they considered foregrounding, defining the task to quantify foregrounding difficult. Thus, we selected these texts because 1) they deal with a similar topic that invites readers to empathize with the suffering of the main character, i.e., a young female protagonist facing a difficult period of her life and manifesting depressive behaviors; and 2) they were rated as examples of good literature, which is functional to the aim of the present study (i.e., of observing differences in readers’ depth of processing of stylistic devices in literary texts). Moreover, both texts make use of internal focalization (Genette, 1983) with a first-person narrator: previous research revealed that first-person narratives (compared to third-person narratives) elicit the highest levels of experience-taking with fictional characters (i.e., the reader’s ability to identify with the character, mimicking her/his personality and inner states; see Creer et al., 2019; Kaufman & Libby, 2012), which may be an important precondition for empathy.

The rationale for using two texts written in distinct literary styles and published in different time periods was to investigate whether the possible relations between empathy and the Processing of Foregrounding would occur independently of those differences. In terms of style, *The Bell Jar* is rich in conventional and elaborate rhetorical elements, employing highly metaphorical language. On the other hand, Lucia Berlin’s short story presents the highly realistic style typical of autofiction (Ellis, 2022), repeatedly employing free indirect discourse and basing its literary quality on the quick flow of images and thoughts that mimics the associations of memory in a stream-of-consciousness.

If all participants were to read the same text, any effects that were to be found could be idiosyncratic to some features present in that particular text. Therefore, randomly assigning participants to read one of the two texts allows us to see whether the relation between the Processing of Foregrounding and empathy can be detected independent of the style and period in which the texts were written. In reporting the results, we took care to check for potential discrepancies between texts and the role of stylistic differences, contextualizing them.

Measures

To measure the Processing of Foregrounding, both quantitative and qualitative measurements were applied in this study through a two-step assessment. First, after having read the full text once, participants were asked to identify through a highlighting task “the parts/wording that you find in the writing style unexpected, unfamiliar, different or disruptive; focus on how it is written (style) and not on what is written (content).” The instruction was adapted from Fialho et al. (2012) and intended to focus readers on the stylistic text features. Second, right after completing the task, participants were asked to respond to an open question: “Describe any thoughts, feelings, images, impressions or memories that were part of your experience related to the parts you have highlighted in the text.” The aim of this two-step measurement was to simplify the original procedure presented by Harash (2021), using self-identified passages instead of eye-tracking data, and an open question instead of the retrospective thinking-aloud protocol.

Empathic reactions were measured by a 30-item, 5-point Likert scale developed by Everson and colleagues (under review; Levett-Jones et al., 2017): the Comprehensive State Empathy Scale (CSES). It consists of six subscales: Empathic Concern; Distress; Shared Affect; Empathic Imagination; Helping Motivation; and Cognitive Empathy. The scale was adapted for literary texts: from the original version by Levett-Jones et al., referring to specific characters’ stories (e.g., “rate the extent to which you experienced each of these feelings in response to the character’s story”) to the adapted version used in the present study, generalizable to different narratives (e.g., “rate the extent to which you yourself experienced each of these feelings in response to the story you read”). Cronbach’s alpha for internal consistency was calculated and considered very high both in the original ($\alpha = 0.95$; Levett-Jones et al., 2017) and in the adaptation for the present study ($\alpha = 0.92$).

The appreciation of the text was assessed through a 4-item 7-point Likert scale. Participants were asked to rate from 1 (completely disagree) to 7 (completely agree) to which degree they agreed on the following statements (from Dixon et al., 1993): “I think this is an example of good literature”; “I enjoyed reading this text”; “I would recommend this text to someone else to read”; “I would be interested to read the rest of this story.”

Finally, participants’ reading expertise was assessed with an open question about the number of books read in the last 12 months.

Coding of open questions

All 78 participants responded to the open question. Their answers were analyzed using the following coding scheme that we based on the work of Harash (2021).

Shallow processing

Readers’ responses that were assigned to this category referred to something strange and unexpected in the text. However, the responses did not report any particular meaning attached to this stylistic element. In their elaborations, readers seemed not sure how to connect such an element to the rest of the story; they seemed not to have spent much time reflecting on it, because they just reported the presence of such an element. This response category can be associated with both positive and negative aesthetic appreciation.

Here is an example from our data set: “Old fashioned colloquial sayings.” The elaboration is synthetic and superficial, referring only generally to the style adopted in the text. No deeper reflection is connected to it.

Failed processing

While reading, readers noticed something in the text that was strange and unexpected to them. In their elaborations, they expressed confusion because of the difficulty in understanding its meaning or the

connection to the rest of the story. Responses that were assigned to this category expressed confusion, irritation, indifference, and negative appreciation.

A representative example from this category would be the following: “I didn’t really empathize with the protagonist and found the wording disjointed: numb trolley bus . . . I’m not sure I understand what that means in this context.” The participant reports having noticed a stylistic distortion which led to confusion and lack of understanding.

Partial processing

While reading, readers noticed something in the text that is strange and unexpected. In their elaborations, they reported the element as striking because it elicited deeper reflections on the underlying meanings of the story. However, readers included in this category were not able to clearly elaborate the deeper insight into the story without arriving at a conclusion, a result, or a solution. In their elaborations, they may express a positive aesthetic appreciation.

An example would be “I found the unusual ways of expressing certain things really made the story. They helped me understand the character and be able to relate what she was experiencing, but not really such that I could describe it clearly to somebody else.” In this elaboration, stylistic distortions are seen as a way to better understand the character’s experience. The reader is still unsure, though, on how to explicitly describe the insight of the story.

Full processing

In their elaborations, readers reported noticing something in the text that was strange and unexpected. They reported the element as striking because it elicited deeper reflections on the underlying meanings of the story. Readers included in this category arrived at a conclusion/result/solution which brought a deeper insight into the story connected, for example, to a symbolic meaning. Overall, readers reported a positive aesthetic appreciation.

An example is “I think the parts I highlighted really show the exclusion from society that the main character feels. The imagery invokes images of death and suffering which create thoughts of othering and survival in my mind. I can really relate to the mind wandering on to morbid topics to try and get away from the realities of the life you are living.” The participants connected the use of stylistic distortions with a deeper message present in the story, the theme of exclusion. Furthermore, in the elaboration, there are clear references to which parts contribute to such insight.

Not applicable

Not all comments could be classified using this coding scheme, predominantly because some readers did not elaborate on their answers enough to enable the annotators to typify their level of processing.

Example: “Immediately thought this was set in Ireland not the USA, but saying that probably during the Vietnam war.” No reference here was made to style and possible interpretations; the elaboration only reports an observation about possible settings of the story.

Two independent annotators assigned participants to one of the five categories (Full, Partial, Failed, Shallow Processing, and Not Applicable). The inter-rater agreement (Krippendorff’s alpha for ordinal variables) was low ($\alpha = .59$ on a range from 0 to 1), where generally, $\alpha \geq .667$ is considered the lowest acceptable result (Krippendorff, 2004). This low agreement was probably due to the short length of participants’ elaborations or the complexity of the judgment required to the annotators. Therefore, each case of disagreement was discussed between the two raters, and a shared agreement was found (see “negotiated agreement,” Campbell et al., 2013, p. 305). To test the validity and reproducibility of the coding proposed, a third independent annotator categorized all responses into the five categories. The agreement between the third annotator and the coding agreed among the first two raters was high, $\alpha = .85$.

Table 1. Ratings divided per category of the three annotators and generally agreed ratings.

Annotators	Foregrounding Processing Categories				
	Full	Partial	Failed	Shallow	Not Applicable
Annotator 1	20	8	12	27	11
Annotator 2	14	16	10	33	5
Annotator 3	18	19	10	22	9
Agreed rating	20	8	9	29	12

Taking this into account, in the following analyses, only the agreed rating will be considered, excluding participants in the Not Applicable category (see, Table 1).

Stylistic analysis

A stylistic analysis was performed by an expert in literary studies. First, the stylistic analysis provides a brief and contextualized overview of the parts that were most often highlighted by participants in the overall sample. Following the procedure explained in Van den Hoven et al. (2016), each word in each text had a score between zero and the total number of readers for each story, that is, 31 for Berlin's and 35 for Plath's text. A score of 0 indicates that the word was never highlighted, and thus, a score of 31 or 35 indicates that the word was highlighted by every participant. In line with previous research (e.g., Kuiken et al., 2004), we considered as a threshold for the following analysis the words that were highlighted by 10 or more participants, which corresponds to approximately 30% of the sample. Second, a comparison between the four groups of readers divided according to their level of Foregrounding Processing is provided, considering the words that were highlighted by at least 30% of the participants. Throughout this section, all stylistic analyses were carried out while constantly referring to the participants' open description of why they found such words and sentences to be striking, thus grounding the theoretical investigation of the identified rhetorical devices within the accuracy of an explicit reader response.

Results

No gender, $\chi^2(1) = 0.00$, $p = .96$, or age, $t(76) = 0.13$, $p = .90$, differences were found between the participants assigned to the two texts. The two texts did not present any differences in key variables, neither in the empathic reactions they elicited in participants, $t(75.06) = 1.31$, $p = .20$, nor in the distributions of Foregrounding Processing categories, $\chi^2(3) = 3.04$, $p = .39$. The data analysis code is openly available in the link reported in the Supplementary Materials.

The mean and standard deviation of the Comprehensive State Empathy Scale (CSES) for each group, excluding the 12 participants categorized as Not Applicable, were calculated and are reported in Table 2.

Exploring the data, the four groups of participants considered in the following analysis (Full, Partial, Failed, Shallow Processing) did not differ for gender, $\chi^2(3) = 1.82$, $p = .61$, or age, $F(3,62) =$

Table 2. Means (M) and standard deviations (SD) of empathy score (from the comprehensive state empathy scale), appreciation, and reading expertise across the four categories.

Variable	Foregrounding Processing Categories							
	Full		Partial		Failed		Shallow	
	M	SD	M	SD	M	SD	M	SD
Empathy score	3.54	0.40	3.74	0.40	2.78	0.77	3.32	0.49
Appreciation	6.19	0.81	5.81	1.22	4.14	2.25	5.39	1.41
Reading expertise	25.58	21.03	35.25	23.16	17.00	18.20	19.26	16.02

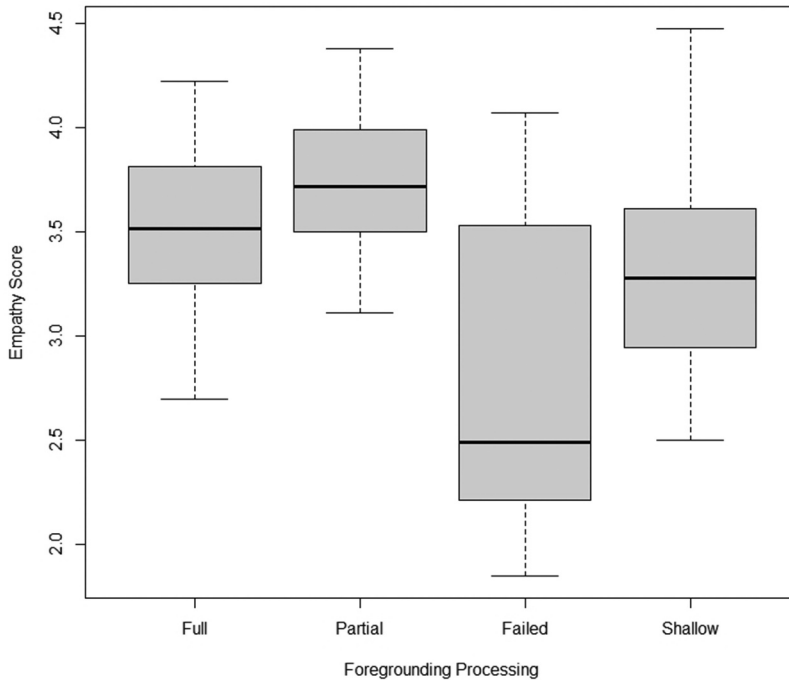


Figure 2. The distribution of the empathic reactions of readers across the four categories of processing of stylistic deviations (i.e., foregrounding).

0.63, $p = .60$. Each group of processing of foregrounding was normally distributed, and the Levene's test showed homogeneity of variance, $F(3,62) = 1.87$, $p = .14$. A one-way independent analysis of variance (ANOVA) was conducted. The distribution of each category and the rating on the CSES score can be seen in [Figure 2](#).

Results showed significant differences in the CSES scores between the groups of Processing of Foregrounding, $F(3,62) = 6.39$, $p < .001$, $\eta^2 = .24$. A Welch's F -ratio one-way test was performed to check whether the results were not biased by the unequal and small sample sizes of the four categories (Field et al., 2012). Results confirmed the ANOVA to be robust, $F(3,20.36) = 4.36$, $p = .02$. Planned contrasts revealed that participants who performed Full, $t(62) = -3.75$, $p < .001$ (one-tailed; see, Field et al., 2012), Partial, $t(62) = 3.92$, $p < .001$ (one-tailed), and Shallow Processing of stylistic deviations, $t(62) = 2.81$, $p < .01$ (one-tailed), had a significantly higher empathic reaction compared to those who performed Failed processing. These results partially align with our hypothesis that deeper processing of Foregrounding (here, Full and Partial) is associated with higher empathic reactions, in this case, when compared with Failed processing. Even though [Figure 2](#) shows that mean empathy is higher for deeper processing of Foregrounding and decreases with less depth of processing, no significant differences were found between the Full and the Partial Processing groups, $t(62) = 0.95$, $p = .35$ (one-tailed); nor between the Full and the Shallow Processing groups, $t(62) = -1.49$, $p = .14$ (one-tailed). The unexpected, relatively high scores in the Partial and Shallow Processing groups and the lack of differences with the Full Processing group might be explained by other factors, namely Appreciation and Reading Expertise, analyzed in the next section.

Contextualizing depth of processing: appreciation and reading expertise

To assess the appreciation of the text, the mean of the four appreciation questions was considered. Overall, the text by Sylvia Plath ($M = 5.85$, $SD = 1.42$) was rated significantly higher than the text of Lucia Berlin ($M = 4.94$, $SD = 1.70$), $t(74.92) = 2.57$, $p = .01$.

The difference in appreciation across the four categories of the depth of Foregrounding Processing was examined (for Means and Standard Deviations, see, Table 2). Exploration of the data showed that the homogeneity of variance across groups was not met, so a Welch's F was tested, and a significant difference across groups was found, $F(3, 19.27) = 3.69, p = .03$. A post-hoc Bonferroni test revealed a significant difference in appreciation between the Full ($M = 6.19, SD = 0.81$) and Failed Processing groups ($M = 4.14, SD = 2.25; p = .003$). No significant differences were found between the other groups, but the Partial and Shallow Processing groups scored higher on appreciation than the Failed Processing group. These results align with the theoretical framework and results reported by Harash (2021): the depth of processing of stylistic distortions is positively related to the appreciation of the literary text. However, it is not yet possible to identify the direction of such a relationship since we do not know, for example, if the participants performed Full Processing of Foregrounding because they appreciated the text more, or vice versa.

With regard to reading expertise, when comparing the four groups of Foregrounding Processing (for means and standard deviations, see Table 2), results showed no significant differences, $F(3, 59) = 1.9, p = .14, \eta^2 = .088$. Even if no significant difference was found, it is worth observing that the Full and Partial Processing groups are higher in their reading expertise compared to Failed and Shallow Processing groups. These results are in line with the framework of Failed Foregrounding (Harash, 2021), which maintains that experienced readers achieve deeper processing of stylistic devices more often compared to inexperienced ones.

Stylistic analysis of the underlinings

To explore whether there are common patterns in what participants underlined in the texts and how such patterns might help to refine our theorizing about a possible relation between foregrounding and empathy, we conducted a stylistic analysis on those text parts that seemed to have stood out to many readers. For this purpose, a literary scholar (the second author) identified and analyzed the rhetorical devices present in the participants' underlinings while grounding theoretical speculations on their possible impact on readers by referring to the self-reported experience of readers themselves, who described why they found these particular combinations of words to be striking. Because the sample size is limited, the present analysis does not aim to state the devices' effect but to provide an in-depth reflection on the relationship between the reader and the text observed in this particular study.

On average, participants highlighted 52.9 words which is 8.94% of Plath's text, and an average of 57.9 words which is 6.13% of Berlin's text. The difference between stories in the number of words highlighted was not significant, $t(64) = 0.36, p = .72$. Furthermore, a high autocorrelation was found between the numbers of highlightings of a word w_i and the following words in both texts (Figure 3), supporting the findings of Van den Hoven et al. (2016): Participants tend to highlight chunks of words (i.e., sentences) instead of single words.

Stylistic analysis of the underlinings: The Bell Jar

In the text of Sylvia Plath, 30 words (belonging to 6 different excerpts) reached the threshold. To facilitate a contextualized understanding, the words will be reported in bold, together with other adjacent words (which, although not reaching the threshold, were often also underlined by participants):

1. *It was a **queer**, sultry summer, the summer they **electrocuted** the Rosenbergs*

The opening sentence of Plath's novel is dense in alliteration ("r" and "s"), outlining the story's time-coordinates with intrinsic thematic allusions. The recounted summer is described as sultry and strange, and its emerging negative connotations are immediately reinforced by an association with death ("electrocuted"). It is possible that participants also found the term "queer" striking as out of

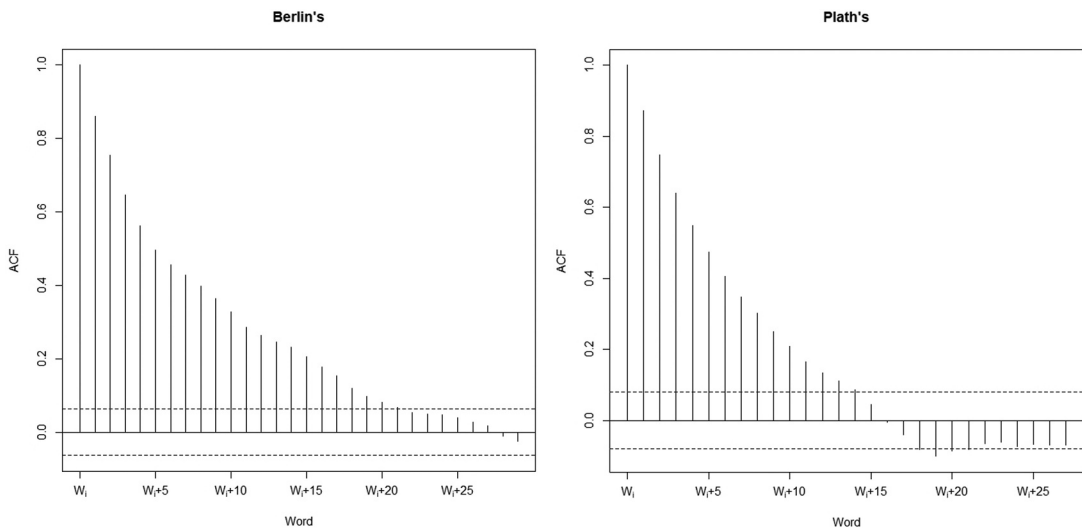


Figure 3. Autocorrelation function (ACF) between the number of highlightings of a word W_i and the following words. It is possible to see how while the lag increases, the autocorrelation decreases, showing how participants tend to highlight chunks of words and not single words.

context, as today the word is predominantly used in LGBTQ+ discourse (e.g., from one participant's report: "I thought that the story would be about the LGBTQ community when I saw the word 'queer'").

2. *I'm stupid about executions*

While the character starts to elaborate on her inner thoughts, this self-directed dysphemism ("stupid") contrasts with its reference to the topic of executions (i.e., there is no clearly identifiable counterpart as being "smart about executions"). The sentence is succinct and sharp; in the words of one participant, it "sounds a bit strange as a stand-alone sentence."

3. *goggle-eyed headlines staring up at me on every street corner and at the fusty, peanut-smelling mouth of every subway*

The news about the execution haunts the protagonist as a recurring personification ("google-eyed headlines"), and the whole city is pervaded by a sensory presence of death: the underground-leading entrances of subways have mouths that are nauseatingly "fusty" and "peanut-smelling."

4. *the fake, country-wet freshness*

The unusual expression ("country-wet freshness") is antithetically defined as "fake," thus immediately unmasking the positive feeling as an early morning illusion, quickly swept away by the dry grayness of the city.

5. *It was like the first time I saw a cadaver*

In a climax, the imagery of death indirectly evoked in the first paragraphs materializes unequivocally as part of the characters' personal experience in another sharp short sentence.

6. *the cadaver's head – or what there was left of it – floated up behind my eggs and bacon*

In describing the haunting feeling that followed the protagonist's sight of a cadaver's head (shown to her by not better identified "Buddy Wilhard"), Plath uses a particularly effective juxtaposition between the macabre image of a cadaver's head and its floating appearance at breakfast, intoxicating the protagonist's daily life routine through the primary and visceral channel of food ("eggs and bacon"). This multi-sensory imagery can likely evoke a sense of nausea, as confirmed by one participant: "[on death] likening it or in the same paragraph as bacon and eggs made my stomach turn."

7. *I felt as though I were carrying the cadaver's head around with me on a string, like some black, noseless balloon stinking of vinegar*

Simile, displaying another morbid juxtaposition between feeling stalked by the memory of the cadaver's head as physically carrying it on a string, evoking the childish image of holding a balloon. In another sensory-charged imagery of death, the balloon is "black," "noseless," and has a fermenting/decomposing smell of "vinegar."

8. *in the company of several anonymous young men with all-American bone structures hired or loaned for the occasion*

The protagonist describes the shallowness of her mundane life in the city, indirectly portrayed as fake and hollow: the young men are described almost as extras on a movie set, "hired" for their ordinary appearance (sarcastically described as "all-American bone structures").

9. *I felt very still and very empty, the way the eye of a tornado must feel*

Simile with a personified "eye of a tornado" that is assumed to have human feelings. This association with the stillness of the eye of a tornado, surrounded by its fierce and devastating winds, is a particularly effective depiction of the numbness of the protagonist, who feels empty and powerless in her inner turmoil, surrounded by the grayness of her busy life in New York.

A representation of the complete highlightings of all participants is displayed in [Figure 4](#).

Although we observed an expected level of variance within highlighted words that do not reach the 30% threshold (Krippendorff's $\alpha = 0.037$), readers, in general, agreed in identifying the most striking stylistic features. In Plath's text, they correspond to effective examples of sensory imagery (particularly targeting the sense of smell) presenting macabre tones and often functioning, potentially, as clues to the protagonist's mental distress. Indeed, many of the short extracts reported above employ rhetorical devices (personification, simile) that serve to juxtapose highly contrasting semantic and sensory domains. Many of these examples of Plath's uncanny and dense metaphorical language (Coyle, 1984) were considered striking by most participants, although the different levels of processing that they performed changed their perception. For example, the same segment was perceived as an obstacle for interpretation by a participant in the Failed Processing group ("The fusty peanut-smelling mouth I was confused by and couldn't picture this image in my mind") and as a functional imaginative tool by a participant in the Full Processing group ("The descriptions are very memorable, evocative of things that we might know, but had not considered. I understood the fusty peanut smell, even though it is not something that I would use to describe myself").

We also investigated more fine-grained differences between the four groups of readers divided according to their level of processing. Both the list of words that were highlighted by at least 30% of participants and the text plot displaying the complete highlighting patterns can be found in Supplementary Materials. For example, readers that experienced Full Processing highlighted very precisely more rhetorical figures compared to the other groups, such as the repetition linking the last two paragraphs ("steering New York like her own private car./Only I wasn't steering anything") leading to the protagonist's self-association with "a numb trolley-bus." The Partial Processing group

It was a **queer, sultry** summer, the summer they **electrocuted** the Rosenbergs, and I didn't know what I was doing in New York: I'm **stupid** about executions. The idea of being electrocuted makes me sick, and that's all there was to read about in the papers: **goggle-eyed** headlines staring up at me on every street corner and at the **fusty, peanut-smelling** mouth of every subway. It had nothing to do with me, but I couldn't help wondering what it would be like, **being burned alive along or your nerves**. I thought it must be the worst thing in the world. New York was bad enough. By nine in the morning the **fake, country-wet** freshness that somehow seeped in overnight evaporated like the tail end of a sweet dream. **Mirage-grey** at the bottom of their **granite** canyons, the hot streets wavered in the sun, the car tops sizzled and glittered, and the dry, cindery dust blew into my eyes and down my throat. I kept hearing about the Rosenbergs over the radio and at the office till I couldn't get them out of my mind. It was like the first time I saw a **cadaver**. For weeks afterwards, the **cadaver's head - or what there was left of it - floated up behind my eggs and bacon at breakfast** and behind the face of Buddy Willard, who was responsible for my seeing it in the first place, and pretty soon I felt as though I were carrying the **cadaver's head around with me on a string, like some black, noseless balloon stinking of vinegar**. I knew something was wrong with me that summer, because all I could think about was the Rosenbergs and how stupid I'd been to buy all those uncomfortable, expensive clothes, **hanging limp as fish in my closet**, and how all the little successes I'd totted up so happily at college **fizzled to nothing** outside the slick marble and plate-glass fronts along Madison Avenue. I was supposed to be having the time of my life. I was supposed to be the envy of thousands of other college girls just like me all over America who wanted nothing more than to be tripping about in those cars, size seven patent leather shoes I'd bought in Bloomington's one lunch hour with a black patent leather belt and black patent leather pocketbook to match. And when my picture came out in the magazine the result of us team working on **drinking martinis in a skimpy, imitation silver-lame bodice stuck on to a big, fat cloud of white tulle**, on some Starlight Roof, in the company of several **anonymous young men with all-American bone structures hired or loaned for the occasion** - everybody would think I must be having a real **whirl**. Look what can happen in this country, they'd say. A girl lives in some **out-of-the-way** town for sixteen years, so poor she can't afford a magazine, and then she gets a scholarship to college and wins a prize here and a prize there and ends up **steering New York like her own private car. Only I wasn't steering anything, not even myself**. I just **bumped** from my hotel to work and to parties and from parties to my hotel and back to work **like a numb trolley-bus**. I guess I should have been excited the way most of the other girls were, but I couldn't get myself to react. I felt very still and very empty, **the way the eye of a tornado must feel, moving dully along in the middle of the surrounding hullabaloo**.

Figure 4. Graphical representation of readers' highlighted words for the extract from Sylvia Plath's *The Bell Jar*. The intensity of the highlighting color, together with the font size and color change accordingly to the number of readers who highlighted the word.

focused more on striking stylistic features in the two paragraphs that are denser in death-related imagery (the first and the third); the Failed Processing group considered less elaborated metaphors ("big, fat cloud of white tulle") and similes ("clothes, hanging limp as fish") as examples of striking elements, and the Shallow Processing groups only agreed on a few occasions, otherwise showing more dispersive patterns in their highlighting.

Stylistic analysis of the underlinings: Stars and Saints

In the text of Lucia Berlin, nine words (belonging to five different passages) reached the threshold. Again, to facilitate a contextualized understanding, the words will be reported in bold, together with other adjacent words in italics.

10. *I did everything to please her, carefully scrolling A.M.D.G. (Ad maiorem Dei gloriam) at the top of every paper*

Latin acronym ("For the Greater Glory of God"), a Jesuit motto employed in Christian schools, which was unfamiliar to many readers.

11. *When I stood up to answer in class they would whisper **Pet, pet, pet***

Repetition of a derogatory remark ("pet") toward the protagonist, reported in free indirect discourse.

12. *It's because I write to him more. **No, you're his Pet.***

The response of the protagonist's mother reiterates her classmate's insult "Pet," capitalized for emphasis and resounding like a final judgment. This unpleasant remark immediately follows the

protagonist's feeble attempt to defend herself from the unreasonable accusation of receiving letters from her dad more often than her mother. The back and forth is succinct and reported in free indirect discourse.

13. *When I was little I didn't see the match, thought she lit her cigarettes with a flaming thumb*

The sight of her mother lighting a match to burn the unopened letter that the protagonist received from her father prompts a highly vivid childhood memory: the simple act of lighting a cigarette metamorphoses in the uncanny image of a "flaming thumb," thus figuratively evoking the portrayal of an abusive mother.

14. *I didn't say, Well now I'm not going to talk anymore*

The protagonist's unspoken decision is again reported in free indirect discourse. The use of the colloquial "well" mimics spoken language, also omitting a subsequent comma that readers expect to see in a written text.

A representation of the complete highlightings of all participants is displayed in Figure 5.

In the overall sample, the number of words highlighted by at least 30% of participants is much lower compared to Sylvia Plath's text. This difference is likely due to the fact that *The Bell Jar* is richer in conventional rhetorical elements, while Lucia Berlin's story presents a realistic, "drier" style strongly based on the use of free indirect discourse. The use of colloquial language, repetitions, and lack of speech marks to signal conversation were the most noticed stylistic distortions, together with the strong visual imagery of the "flaming thumb." This writing style was sometimes found to be obstructive, as expressed by a participant from the Partial Processing group: "I did find the piece

Whenever a siren sounded outside in the streets, **near or far**, Sister Cecilia had us stop whatever we were doing, lay our heads down on our desks, and say a Hail Mary. **I still do that. Say a Hail Mary, I mean. Well, also I lay my head down on wooden desks, to listen to them, because they do make sounds, like branches in the wind, as if they were still trees.** A lot of things were really bothering me in those days, like what gave life to the candles and where the sound came from in the desks. If everything in God's world has a soul, even the desks, **since they have a voice, there must be a heaven. I couldn't go to heaven because I was Protestant.** I'd have to go to limbo. I would rather have gone to hell than limbo, what an ugly word, like **dumbo, or mumbo jumbo**, a place without any dignity at all. I told my mother I wanted to become a Catholic. She and my grandpa had a fit. He wanted to put me back in Vitas school but she said no, it was full of Mexicans and juvenile delinquents. I told her there were lots of Mexicans at St. Joseph's but she said they came from nice families. **Were we a nice family? I didn't know. What I still do is look in picture windows where families are sitting around and wonder what they do, how do they talk to one another?** Sister Cecilia and another nun came to our house one afternoon. I don't know why they came and they didn't get a chance to say: **Everything was a mess. My mother crying and Mamie, my grandma, crying. Grandpa was drunk and went lunging at them calling them crows.** The next day I was afraid Sister Cecilia would be mad at me and not say "Goodbye, dear" when she left me alone in the room at recess, the before she left she handed me a book called **Understood Betty** and said she thought I would like it. It was the first real book I ever read, the first book I fell in love with. **She prized my work in class, and commented to the other students every time I got a star, or on Fridays when I was given a Saint card.** I did everything to please her, carefully **scrolling A.M.D.G. (Ad maiorem Dei gloriam)** at the top of every paper, rushing to erase the board. My prayers were the loudest, my hand the first to go up when she asked a question. She continued to give me books to read and **once she gave me a paper bookmark that said "Pray for us sinners now and at the hour of our death."** I showed it to Melissa Barnes in the cafeteria. I had foolishly believed that since Sister Cecilia liked me the girls would begin to like me too. But now instead of **laughing at me they hated me.** When I stood up to answer in class they would whisper **Pet, pet, pet.** Sister Cecilia chose me to collect the dimes and pass out the medals for lunch and **when each girl took her medal she whispered Pet.** Then one day out of the **clear blue sky** my mother got mad at me because my father **wrote me more than he did her. It's because I write to him more. No, you're his Pet.** One day I got home late. I had missed the bus from the plaza. She stood at the top of the stairs with a blue airmail letter from my father in one hand. With the other she lit a kitchen match on her thumbnail and burned the letter as I raced up the stairs. That always scared me. When I was little I didn't see the match, **thought she lit her cigarettes with a flaming thumb.** I stopped talking. **I didn't say, Well now I'm not going to talk anymore.** I just gradually stopped and when the sirens passed I laid my head down on the desk and whispered the prayer to myself. When Sister Cecilia called on me I shook my head and sat back down. I stopped getting Saints and stars. It was too late. Now they called me **dumb-dumb.** She stayed in the classroom after they had left for gym. "What is wrong, dear? May I help you? Please talk to me." I locked my jaws and refused to look at her. She left and I sat there in the hot **semidarkness** of the classroom. She came back, later, with a copy of *Black Beauty* that she placed before me. "This is a lovely book, only it's very sad. Tell me, are you sad about something?" I ran away from her and the book into the cloakroom. **Of course there were no cloaks since it was so hot in Texas,** but boxes of dusty textbooks, Easter decorations, Christmas decorations. Sister Cecilia followed me into the tiny room. She spun me around and **forced me to my knees.** "Let us pray," she said. Hail Mary, full of grace, the Lord is with thee. Blessed is the fruit of thy world, Jesus... Her eyes were filled with tears. I could not bear their tenderness. I wrenched away from her grasp, accidentally knocking her down. Her headdress caught in a coat hook and was **yanked off. Her head wasn't shaved like the girls said.** She cried out and ran from the room. I was sent home that same day, expelled from St Joseph's for **striking a nun.** I don't know how she could have thought that I would hit her. **It wasn't like that at all.**

Figure 5. Graphical representation of readers' highlighted words for the extract from Lucia Berlin's *Stars and Saints*. The intensity of the highlighting color, together with the font size and color change accordingly to the number of readers who highlighted the word.

confusing at times, particularly with time switching back and forth and with failure to note that someone was speaking.” Other times, it was appreciated as a stylistic device that showed the thinking pattern of the main character: “It was almost as though we were inside her brain and processing her thoughts with her” (participant from the Full Processing group).

As with the previous excerpt, more fine-grained reports divided by the four categories of processing (words that reached the threshold and plots of highlighting patterns) are reported in the Supplementary Materials. Overall, participants in the Full Processing group were found again to highlight more often (and more precisely) striking stylistic features that went beyond the free indirect speech, such as the assonance “mumbo jumbo” and the auditory imagery in the simile of the protagonist “listening” to the wooden desks “because they do make sounds, like branches in the wind, as if they were still trees.” The Partial Processing group also agreed on slightly more stylistic deviations compared to the overall sample, while both the Failed and Shallow Processing groups agreed only on a couple of examples of colloquial indirect speech, showing even more variance in the other highlighted portions of the text. However, we already noted that this tendency is more general: Berlin’s writing style in this excerpt elicited low agreement on the expressions that readers found striking (Krippendorff’s $\alpha = 0.052$), although this did not lead to an overall significantly lower percentage of highlighted words than Plath’s; thus, subjective experience likely played a more prominent role in selecting which words to highlight than it did with Plath’s text.

Discussion

Results support the hypothesis that readers’ processing of foregrounding plays a role in how literary reading can affect empathic reactions. The outcome of the present study partially supports our prediction that a higher depth in processing stylistic distortions is associated with stronger empathic reactions in readers. In fact, participants who performed Full and Partial processing reported significantly higher empathic reactions than participants who failed to process stylistic distortions. However, contrary to our expectations, no significant difference was found between the Full and Shallow Processing groups. In addition, no difference was found between the Full and Partial Processing groups. In fact, the Partial Processing group had a higher score on empathic reactions compared to the Full Processing group. Even though this difference was not significant, this unexpected result might suggest that it is not only the full processing of stylistic distortions that can elicit the highest empathic reactions from readers. The group that fully processed foregrounding showed higher aesthetic appreciation of the text than each of the other groups and significantly higher than the Failed Processing group. Instead, the other groups did not differ in appreciation from each other. Furthermore, both the Full and Partial processing groups were higher in reading experience (even though not significant) than the other two groups. Therefore, it may be that not only the Full Processing of stylistic distortions but also an incomplete resolution of the meaning gap, together with a positive aesthetic appreciation of the text, create the highest empathic reaction or, as stated by Shklovsky (1917/2016, p. 80), “the process of perception is its own end in art.”

The differences between the groups of Foregrounding Processing also emerged in the analysis of the text parts that participants underlined. The Shallow Processing of the texts seemed associated with a rather dispersive pattern of underlinings, revealing no unanimity about what stands out in the text. In contrast, full foregrounding seemed more targeted in the underlining task, focusing on style aspects well-known in stylistics and that we can hypothesize to be helpful in providing readers access to the character’s state of mind. It is important to note that those very same aspects could be experienced as obstacles to an understanding (and appreciation), in particular in the Failed Foregrounding group. Taking into consideration the results so far, we suggest that the devices that were the focus of the Full Foregrounding group may contribute to a fuller experience (e.g., through imagery) of the story, which may be dovetailing with a fuller empathic response to characters. In contrast, the Failed Foregrounding group marked text parts that seem much less effective in understanding the mind of the characters. Nonetheless, it is not

yet possible to establish which factors are intervening in the relationship between Foregrounding Process and empathy. In fact, it is still an open question whether it is the depth of the processing of stylistic devices that enhances empathic reactions or whether it is a predisposition of the reader, such as trait empathy, that directs the attention toward specific elements in the text which provide access to the state of mind of the character. Further research is needed to answer this question, considering different readers' traits.

Starting from the observation reported by Harash (2021), and as emerged in the present study, one element that distinguishes Failed Processing from the other categories is the negative aesthetic experience of the literary text, probably due to difficulties in understanding and mostly associated with feelings of frustration (see also Bálint et al., 2016). Further investigations are needed to disentangle the direction of the relationship between the appreciation of a literary text and the depth of processing of its stylistic devices. Actually, it is not clear yet if it is the depth of elaboration of the stylistic elements that directs the reader toward a higher understanding of the beauty of the artwork, or vice versa.

The trend in aesthetic appreciation is also coherent with the patterns we observed in the stylistic analysis of the textual features perceived as striking and unfamiliar. On the one hand, participants who experienced either full or partial processing – although more prominently in the first group – seemed to agree on a higher number of stylistic elements, and they also tended to include more elaborate rhetorical elements that participants in other groups overlooked. On the other hand, the Shallow and Failed Processing groups agreed on fewer occasions, including more common stylistic elements that were not perceived as striking by the more avid readers of the first two groups. However, on a more general level, we still observed a considerable overlap across all groups for at least some striking stylistic features – which allowed for a fine-grained exploration of the differences in *how* readers process these features and evaluate them on an aesthetic level. Indeed, the same stylistic deviations were described as unpleasant and unresolved disruptions of the interpretative process by readers in the Failed Processing group and as relatively neutral by readers in the Shallow Processing group; whereas readers in the other two groups found such disruptions to be aesthetically pleasing and carrying “deeper meanings,” with the Full Processing group managing to unfold these meanings by providing their personal interpretative resolution. Several participants in the Partial and Full Processing groups reported in their open answers that the stylistic deviations in the texts made them feel closer to the character, functioning as pleasurable “riddles” for understanding the experience. Thus, our qualitative exploration further enriches what we found in the quantitative analysis that connected levels of processing to CSES scores: higher levels in the depth of processing of literary texts, together with a positive aesthetic appreciation of the text, likely goes hand in hand with a higher chance of experiencing empathic reactions. Therefore, the present results suggest empathic reactions might take place not only at the end of the full process (Dehabituation–Reflection–Refamiliarization; Figure 1) but along the reading experience, with positive aesthetic appreciation playing a pivotal, perhaps mediating, role.

It is interesting to note that no significant differences were found between the two texts on the level of empathy they elicited and the distribution of Foregrounding Processing categories, even if initially, the two texts were rated differently in foregrounding level by experts. Although very preliminary, this result might support the idea of avoiding defining literature based on conventions determined by an elite group of scholars and instead adopting a bottom-up approach, focusing on what and how is perceived as literature by the reader.

Finally, it is worth mentioning another possible explanation of the difference between the four processing categories: the readers' expertise, defined here as the amount of exposition to literary texts. In the present results, no significant differences were found between the four groups, even though the Full and Partial Processing groups showed higher levels of reading expertise than the Failed and Shallow Processing groups. The lack of significance in those differences contrasts with what was reported in previous research (Harash, 2021) and might be related to the measure we used to assess reading expertise. Future studies should include a more sensitive operationalization, for example,

using the Author Recognition Test (ART, Stanovich & West, 1989). Such a test was not appropriate in the present design because of potential cultural differences among participants due to the online data collection, but it might be useful with a different design. Or a more fine-grained measure for reading experience that is not culturally sensitive might be developed.

Limitations and future directions

The present study carries some limitations. First, we have to stress that our study design did not include random assignment to conditions of the independent variable of Processing of Foregrounding. Rather, four groups of varying levels of processing of foregrounding were post-hoc created based on the coding of readers' answers. Therefore, no causal direction of effects can be assessed based on the present results; interpretations should be limited to associations between foregrounding and empathy, which cannot exclude the other way around: empathy may also influence the depth of Processing of Foregrounding. Second, the qualitative analysis concerned post-reading reflections, and it is not possible to determine how shallow or deep such a process took place in readers' minds during the *online* processing. Third, as mentioned above, the qualitative analysis of readers' elaborations required a certain level of interpretation by the three coders; this indeed may potentially hinder the reproducibility of the method applied. It is possible, for example, that what was Shallow was not the processing of Foregrounding but the way the readers reported their processing and the depth of their description in the short answer required. This would also possibly explain the unexpected lack of difference between the Shallow and Full Processing groups. In addition, also the stylistic analysis performed on the participants' highlightings required an interpretative move from the researchers. In order to limit possible biases and enhance reproducibility, next studies should consider, for example, having a follow-up interview with participants to clarify their impressions and experiences about the text read. Moreover, in the present exploratory work, we only observed the behavior of readers in two texts. Therefore, it is necessary to investigate the connection between Foregrounding Processing, empathic reactions, and appreciation in other texts as well. This will possibly improve the coding scheme presented here.

A final limitation is due to the restrictions of the coding system and the exploratory nature of the present study, in which Foregrounding Processing was divided into four categories. From a theoretical perspective, it may be more plausible for the Processing to be a continuous measure. Thus, further studies need to focus on developing an assessment tool that can grasp the complexity of this Processing, and our results suggest that such research endeavors will likely benefit from considering the potential roles of appreciation and (probably) reading expertise as well.

Conclusion

The present research examined which aspects of reading literary texts "experience may be responsible for readers" empathic responses toward the story character. With an approach focused on readers' perceptions of stylistic deviations and specific attention to how they processed such deviations, the present study provided preliminary support to the hypothesis that the depth of Processing of Foregrounding is connected with higher empathic reactions. However, results revealed that all Full, Partial, and Shallow Processing groups have relatively high levels of empathy, with the group of Partial processing scoring the highest and the Failed group the lowest. Such an outcome seems to be related to the appreciation of the text that might further influence the reading experience and empathic response. The present study sheds light on the controversial results described in previous research on the impact of literary reading on empathic reactions. In fact, it seems necessary to consider readers as a heterogeneous group widely varying in how they perceive stylistic devices and how they process them. As revealed in this research, the differences in Foregrounding Processing are connected with different empathic reactions, and this relationship, not taken into consideration in previous research, might have confounded the results of experimental manipulations. Indeed, the next steps in the

empirical study of the effects of reading literature need to take into account personal differences in both readers' perception and processing of stylistic features.

Acknowledgments

This work was supported by the VU Amsterdam, the Utrecht University, and the ELIT Network. We thank: all the participants for their precious contribution; Amir Harash for his input on his framework and the present adaptation; Jonathan Cohen for his valuable feedback in the designing process of this study; Kirren Chana for her precious proofreading.

Disclosure statement

No potential conflict of interest was reported by the author(s).


Funding

This work was supported by the European Union's H2020 Marie Skłodowska-Curie Actions Research and Innovation Program No 860516.

ORCID

Giulia Scapin  <http://orcid.org/0000-0002-6163-714X>

Cristina Loi  <http://orcid.org/0000-0002-3694-5837>

Frank Hakemulder  <http://orcid.org/0000-0003-3923-8668>

Katalin Bálint  <http://orcid.org/0000-0002-6254-7883>

Elly Konijn  <http://orcid.org/0000-0002-6134-8700>

Data availability statement

The data that support the findings of this study are openly available in OSF at https://osf.io/mu5ad/?view_only=9206294177f84938bd8542b3fe638935.

References

- Ahearne, Jeremy, & Speller, John (Eds.). (2012). *Bourdieu and the Literary Field* (Vol. 35). Edinburgh University Press. <https://doi.org/10.1515/9781474463829>
- Bálint, K., Hakemulder, F., Kuijpers, M. M., Doicaru, M. M., & Tan, E. S. (2016). Reconceptualizing foregrounding. *Scientific Study of Literature*, 6(2), 176–207. <https://doi.org/10.1075/ssol.6.2.02bal>
- Bálint, K., Sukalla, F., & Rooney, B. (2022). Personal relevance and state empathy with a character facilitates self-disclosure in film viewers. *Frontiers in Communication*, 7, 1–18. [984341]. <https://doi.org/10.3389/fcomm.2022.984341>
- Berkers, P. (2009). Ethnic boundaries in national literary histories: Classification of ethnic minority fiction authors in American, Dutch and German anthologies and literary history books, 1978–2006. *Poetics*, 37(5–6), 419–438. <https://doi.org/10.1016/j.poetic.2009.09.003>
- Berlin, L. (2016). *A manual for cleaning women: Selected stories*. Picador USA.
- Brown, D. (2003). *The Da Vinci code*. Random House.
- Calarco, N., Fong, K., Rain, M., & Mar, R. A. (2017). Absorption in narrative fiction and its possible impact on social abilities. In F. Hakemulder, M. M. Kuijpers, & K. E. S. H. Tan (Eds.), *The handbook of narrative absorption* (pp. 293–313). Balint and MM Doicaru.
- Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, 42, 294–320. <https://doi.org/10.1177/0049124113500475>
- Cook, G. (1994). *Discourse and literature: The interplay of form and mind*. Oxford University Press.
- Coyle, S. (1984). Images of madness and retrieval: An exploration of metaphor in The Bell Jar. *Studies in American Fiction*, 12(2), 161–174. <https://doi.org/10.1353/saf.1984.0027>

- Creer, S. D., Cook, A. E., & O'Brien, E. J. (2019). Taking the perspective of the narrator. *Quarterly Journal of Experimental Psychology*, 72(5), 1055–1067. <https://doi.org/10.1177/1747021818779496>
- De Mulder, H. N., Hakemulder, F., Klaassen, F., Junge, C. M., Hoijsink, H., & van Berkum, J. J. (2022). Figuring out what they feel: Exposure to eudaimonic narrative fiction is related to mentalizing ability. *Psychology of Aesthetics, Creativity, and the Arts*, 16(2), 242. <https://doi.org/10.1037/aca0000428>
- Dixon, P., Bortolussi, M., Twilley, L. C., & Leung, A. (1993). Literary processing and interpretation: Towards empirical foundations. *Poetics*, 22(1–2), 5–33. [https://doi.org/10.1016/0304-422X\(93\)90018-C](https://doi.org/10.1016/0304-422X(93)90018-C)
- Dodell-Feder, D., & Tamir, D. (2018). Fiction reading has a small positive impact on social cognition: A meta-analysis. *Journal of Experimental Psychology: General*, 147(11), 1713–1727. <https://doi.org/10.1037/xge0000395.supp>
- Ellis, N. (2022). The short autofictions of Eve Babitz, Lucia Berlin and Bette Howland. *Critique: Studies in Contemporary Fiction*. <https://doi.org/10.1080/00111619.2022.2097049>
- Fialho, O. (2019). What is literature for? The role of transformative reading. *Cogent Arts & Humanities*, 6(1), 1692532. <https://doi.org/10.1080/23311983.2019.1692532>
- Fialho, O., Miall, D. S., & Zyngier, S. (2012). Experiencing or interpreting literature: Wording instructions. *Pedagogical Stylistics: Current Trends in Language, Literature and ELT*, 58, 74.
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage Publications.
- Genette, G. (1983). *Narrative discourse: An essay in method* (Vol. 3). Cornell University Press.
- Hakemulder, F. (2000). *The moral laboratory: Experiments examining the effects of reading literature on social perception and moral self-concept*. Benjamins.
- Hakemulder, F. (2020). Finding meaning through literature: Foregrounding as an emergent effect. *Anglistik: International Journal of English Studies*, 31(1), 91–110. <https://doi.org/10.33675/ANGL/2020/1/8>
- Harash, A. (2019). *Attention, aesthetic appraisal and semantic noise while reading of a literary text*. [Unpublished doctoral dissertation or master's thesis]. Tel Aviv University.
- Harash, A. (2021). The model of failed foregrounding. *Psychology of aesthetics, creativity, and the arts*, 412–425. Advance online publication. <https://doi.org/10.1037/aca0000419>
- Herrnstein Smith, B. (1988). *Contingencies of value: Alternative perspectives for critical theory*. Harvard University Press.
- Janssen, S. (2001). The empirical study of careers in literature and the arts. In D. Schram & G. J. Steen (Eds.), *The psychology and sociology of literature; In honor of Elrud Ibsch* (pp. 323–357). John Benjamins Publishing.
- Johnson, D. R. (2013). Transportation into literary fiction reduces prejudice against and increases empathy for Arab-Muslims. *Scientific Study of Literature*, 3(1), 77–92. <https://doi.org/10.1075/ssol.3.1.08joh>
- Kaufman, G. F., & Libby, L. K. (2012). Changing beliefs and behavior through experience-taking. *Journal of Personality and Social Psychology*, 103(1), 1. <https://doi.org/10.1037/a0027525>
- Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, 342(6156), 377–380. <https://doi.org/10.1126/science.1239918>
- Koek, M. (2022). *Think twice: Literature lessons that matter*. [Unpublished dissertation]. University of Amsterdam.
- Koopman, E. M. (2016). Effects of 'literariness' on emotions and on empathy and reflection after reading. *Psychology of Aesthetics, Creativity, and the Arts*, 10(1), 82–98 doi:10.1037/aca0000041.
- Koopman, E. M. E., & Hakemulder, F. (2015). Effects of literature on empathy and self-reflection: A theoretical-empirical framework. *Journal of Literary Theory*, 9(1), 79–111. <https://doi.org/10.1515/jlt-2015-0005>
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage Publications.
- Kuiken, D., Phillips, L., Gregus, M., Miall, D. S., Verbitsky, M., & Tonkonogy, A. (2004). Locating self-modifying feelings within literary reading. *Discourse Processes*, 38(2), 267–286. https://doi.org/10.1207/s15326950dp3802_6
- Kuzmičová, A., Mangen, A., Støle, H., & Begnum, A. C. (2017). Literature and readers' empathy: A qualitative text manipulation study. *Language and Literature: International Journal of Stylistics*, 26(2), 137–152. <https://doi.org/10.1177/0963947017704729>
- Levett-Jones, T., Lapkin, S., Govind, N., Pich, J., Hoffman, K., Jeong, S. Y., Norton, C. A., Noble, D., Maclellan, L., Robinson-Reilly, M., & Everson, N. (2017). Measuring the impact of a 'point of view' disability simulation on nursing students' empathy using the comprehensive state empathy scale. *Nurse Education Today*, 59, 75–81. <https://doi.org/10.1016/j.nedt.2017.09.007>
- Mar, R. A. (2018). Evaluating whether stories can promote social cognition: Introducing the Social Processes and Content Entrained by Narrative (SPaCEN) framework. *Discourse Processes*, 55(5–6), 454–479. <https://doi.org/10.1080/0163853X.2018.1448209>
- Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communications*, 34(4), 407–428. <https://doi.org/10.1515/COMM.2009.025>
- Miall, D. S. (2006). *Literary reading: Empirical & theoretical studies*. P. Lang.
- Miall, D. S., & Kuiken, D. (1994). Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics*, 22(5), 389–407. [https://doi.org/10.1016/0304-422X\(94\)00011-5](https://doi.org/10.1016/0304-422X(94)00011-5)
- Miall, D. S., & Kuiken, D. (1999). What is literariness? Three components of literary reading. *Discourse Processes*, 28(2), 121–138. <https://doi.org/10.1080/01638539909545076>

- Mumper, M. L., & Gerrig, R. J. (2017). Leisure reading and social cognition: A meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 11(1), 109–120. <https://doi.org/10.1037/aca0000089>
- Oatley, K. (2012). The cognitive science of fiction. Wiley interdisciplinary reviews. *Cognitive Science*, 3(4), 425–430. <https://doi.org/10.1002/wcs.1185>
- Plath, S. (1963/2013). *The Bell Jar*. Faber & Faber.
- Quinlan, J. A., Padgett, J. K., Khajehnasiri, A., & Mar, R. A. (2022). Does a brief exposure to literary fiction improve social ability? Assessing the evidential value of published studies with a p-curve. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0001302>
- Samur, D., Tops, M., & Koole, S. L. (2018). Does a single session of reading literary fiction prime enhanced mentalising performance? Four replication experiments of Kidd and Castano (2013). *Cognition and Emotion*, 32(1), 130–144. <https://doi.org/10.1080/02699931.2017.1279591>
- Scapin, G. (2022). *Book power: Reducing stigmatization of depression through contact with literary characters*. Poster Session organized by the graduate school of the faculty of social science of the Vrije Universiteit Amsterdam. <https://doi.org/10.13140/RG.2.2.13683.40489>
- Shklovsky, V. (1917/2016). Art as device (A. Berlina, Trans.). In A. Berlina (Ed.), *Viktor Shklovsky: A reader* (pp. 73–96). Bloomsbury.
- Simpson, P. (2014). *Stylistics. A resource book for students* (2nd ed.). Routledge.
- Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 24(4), 402–433. <https://doi.org/10.2307/747605>
- Van den Hoven, E., Hartung, F., Burke, M., & Willems, R. M. (2016). Individual differences in sensitivity to style during literary reading: Insights from eye-tracking. *Collabra*, 2(1). <https://doi.org/10.1525/collabra.39>
- Van Kuijk, I., Verhoeven, P., Dijkstra, K., & Zwaan, R. A. (2018). The effect of reading a short passage of literary fiction on theory of mind: A replication of Kidd and Castano (2013). *Collabra: Psychology*, 4(1). <https://doi.org/10.1525/collabra.117>
- Van Peer, W., Sopcak, P., Castiglione, D., Fialho, O., Jacobs, A. M., & Hakemulder, F. (2021). Foregrounding. In D. Kuiken & A. Jacobs (Eds.), *Living handbook of the empirical study of literature* (pp. 145–176). De Gruyter.
- Verboord, M. (2003). Classification of authors by literary prestige. *Poetics*, 31(3–4), 259–281. [https://doi.org/10.1016/S0304-422X\(03\)00037-8](https://doi.org/10.1016/S0304-422X(03)00037-8)
- Wimmer, L., Friend, S., Currie, G., & Ferguson, H. J. (2021). Reading fictional narratives to improve social and moral cognition: The influence of narrative perspective, transportation, and identification. *Frontiers in Communication*, 5, 611935. <https://doi.org/10.3389/fcomm.2020.611935>
- Zwaan, R. A. (1991). Some parameters of literary and news comprehension: Effects of discourse-type perspective on reading rate and surface structure representation. *Poetics*, 20(2), 139–156. [https://doi.org/10.1016/0304-422X\(91\)90003-8](https://doi.org/10.1016/0304-422X(91)90003-8)