Discourse Processes

Understanding and Appreciating Literary Texts Through Rereading

Moniek M. Kuijpers & Frank Hakemulder

Routledge
Taylor & Francis Group

Understanding and Appreciating Literary Texts Through Rereading

Moniek M. Kuijpers [a,b*] and Frank Hakemulder[b]

aDepartment of Language & Literature, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany; bDepartment of Media and Culture, Utrecht University, Utrecht, Netherlands

ABSTRACT

Previous research showed an emerging appreciation of literary narratives on second reading, whereas such effects fail to occur for the same narratives depleted of literary features. This might suggest that appreciation is associated with readers' acknowledgment of the purposefulness of literary devices on rereading. It may also be that the increase in appreciation is caused by a general sense of increased comprehension, a more common effect that may also occur on rereading nonliterary narratives. Three studies were conducted in which participants reread either original literary texts or manipulated versions in which literary style aspects were normalized. Using linear mixed models we examined the relationship between levels of literariness, perceived comprehension, and appreciation as well as the mediating influence of participants' reading experience. The results show that an increase in appreciation seems mainly related to an increase in perceived comprehension, independent of the level of literariness.

Introduction

Texts vary in the rewards they yield to their readers. This holds even more so for rereading them: For some texts one reading suffices, whereas others seem to promise there is more to be discovered on second reading. Consequently, some readers return to the novels they read years ago and find that either they or the book seem to have changed. They may feel the urge to revisit passages earlier on in a story they are reading, and they may feel compelled to read poems repeatedly before moving on to the next. Strangely enough, having prior knowledge of a text does not seem to curtail responses such as transportation (cf. Green et al., 2008), enjoyment (cf. Leavitt & Christenfeld, 2011), and strong aesthetic responses (Wassiliwizky, Koelsch, Wagner, Jacobsen, & Menninghaus, 2017). Qualitative data (Bálint, Hakemulder, Kuijpers, Tan, & Doicaru, 2016) suggest that when readers encounter deviating aspects in a text, they may choose to reread as part of their strategy to prolong their contact with the text, to disambiguate parts they feel uncertain about, or to slow down the pace of their reading as they experience an overpowering force of the narrative style. To the participants this rereading enhances rather than obstructs their engagement with the text.

Voluntary repeated exposure seems common enough among audiences of various genres (e.g., poetry, but also religious and instructional texts) and media (e.g., written texts, but also movies and paintings). However, rereading *literary* texts is considered to be exceptionally rewarding. Indicative of this assumption are the many practices in Western societies (e.g., literary education, reading groups, literary criticism, and literary studies) that aim to mine their riches. The notion of a hidden message *in the way* these texts are formulated is a central claim in literary studies (Hakemulder & Van Peer, 2015). Hence, it is expected to be worthwhile to analyze and reflect on literary texts and, in support of that process, reread them. As such, it may be "literariness" (cf. Jakobson's *literaturnost*;

see Stempel, 1972), or the unique qualities of literary style, that makes the rereading of literature so particularity gratifying. Some literary scholars even argue that a distinctive aspect of literariness seems to be that it typically emerges over time rather than on first reading and that it is the result of an interaction between reader and text (Dixon, Bortolussi, Twilley, & Leung, 1993; Hakemulder, 2004, 2008; Zyngier, van Peer, & Hakemulder, 2007).

Additionally, with multiple exposures to a text increased appreciation is expected to arise, rather than mere improved information processing. Previous work has shown that rereading texts with specific literary features does enhance appreciation as compared with evaluation after a first reading, whereas such an emergent effect fails to occur in their absence (Dixon et al., 1993). It remains unclear, however, what it *is* exactly that emerges. The challenges of reading literary texts may include grasping the significance of the author's style choices. But readers simultaneously need to understand how the events described make up the plot, what the interrelationships between characters are, what their behavioral motives are, estimate the reliability of narrators, and so on. Processing these aspects may benefit from rereading, more or less independent of literary style. To investigate this possibility we designed a series of experiments in which we compared the influence of two factors on increases in appreciation: literariness of the texts and increases in perceived comprehension. Thus, we hoped to establish whether the effects are particular to literature or a more common process that we may see in nonliterary narratives as well.

Literariness

There is a number of reasons to assume that rereading literature is a special case, distinct from rereading nonliterary texts. The results of the study by Dixon et al. (1993) suggest that literariness is an effect that emerges during the interaction between reader and identifiable text qualities. Their participants read one of two versions of a story by Borges (*Emma Zunz*): either the original or one rewritten by the researchers. In the former, readers are confronted with a narrator who is unusually oblivious to aspects of the fictional world, a characteristic that researchers argued to be crucial for readers to fully appreciate the literary value of the story. In their rewritten text the ambiguous narrator was changed into an omniscient one, which seems a more common choice in narrative fiction. After reading one of the versions participants evaluated the story, read it once more, and evaluated it again. As the researchers had expected, after rereading the original text appreciation increased, whereas evaluation of the manipulated version remained unchanged after a second reading. However, this held only for a part of their sample, that is, the frequent readers among the participants. Hence, the emergent effect depends on text qualities but also on who reads.

Dixon et al. (1993) suggested that their rereading paradigm might help to find an empirical basis for literariness, with literariness defined as an effect that emerges as the result of an interplay between specific literary text qualities and their readers. They suggested this effect reveals itself in an increased appreciation, or "depth of appreciation," that can be operationalized as the difference between first and second evaluation.

This notion of literariness was picked up later by other researchers, using different texts but comparable text manipulations. In a study by Hakemulder (2004) participants either reread a poem by Nabokov or a manipulated version in which the language usage had been normalized by the researcher. In a study by Zyngier et al. (2007) participants reread poetry lines of various complexity levels. In a study by Hakemulder (2008) movie adaptations of Shakespeare plays were used. Participants either watched a scene from a mainstream adaptation twice or a corresponding scene from an atypical adaptation. The results of all these studies confirmed the conclusions of Dixon et al. (1993): Deviation from standard representations leads to increased appreciation on second exposure, whereas without such deviations such increases do not occur. These studies do not clarify, however, what the nature of the emergent literary effects might be. It remains unclear, as Dixon et al. emphasized (1993, p. 14), what those literary effects are. For instance, we do not know whether the participants in their study, all nonexpert readers, noted the significance of the ambiguous

narrator, let alone whether they came to the same conclusion as the researchers about its specific relevance for the interpretation of that literary text (assuming *that* is what they were working on during their rereading).

Literary studies, and in particular foregrounding theory, may provide a theoretical framework that helps us conceptualize the nature of literary effects (e.g., Mukarovsky, 1964; Van Peer, 1986). It suggests that the use of language in literature typically deviates from daily usage. These deviations violate certain norms or conventions (e.g., grammar rules) or break with a pattern set within the text itself (e.g., an irregularity in a poem's rhyme scheme). They are assumed to draw readers' attention, slow down text processing, generate new insights or renew awareness (i.e., deautomatization, or *Verfremdung*; Brecht, 1976; Shklovsky, 1965). In other words, on top of communicating the basic facts about narrative events, foregrounding may result in an additional layer of information (e.g., a metaphor conveying insights about these events or causing some effect adding to the overall point of the story). In each of the studies mentioned above it is some deviating aspect (e.g., the ambiguous narrator, low-frequency words, unusual camera angles) that researchers suggested are responsible for the literary effect that arises in the interaction between reader and these deviating text elements.

So far a number of empirical studies have investigated the effects of the use of foregrounding (Hakemulder & Van Peer, 2015), some of which did indeed show a relation between deviating text qualities and appreciation. These studies mainly focused on the effect described as defamiliarization within foregrounding theory. Miall and Kuiken (1994), for example, found that foregrounding in a text captures readers' attention and leads readers to evaluate a text as more striking and evocative (cf. Hunt & Vipond, 1985; Van Peer, 1986). Emmott and colleagues (Sanford & Emmott, 2012; Sanford, Sanford, Molle, & Emmott, 2006) showed how literary devices capture readers' attention. Although most language is often processed in a fairly shallow way, deviations stimulate a *deeper* processing, which might be experienced as more rewarding. Similar conclusions can be drawn from research on cognitive disfluency (Alter, 2013; Menninghaus et al., 2015). These studies suggest that "cognitive roadblocks" (Alter, 2013, p. 237) signal the necessity for deeper processing and would hence improve understanding and learning results, stimulating generalizing from concrete examples, concentrating more on global aspects rather than sticking to narrower features. Overcoming moderate difficulties and reobtaining processing fluency, and the satisfaction of "getting it" and experiencing agency (cf. Bálint et al., 2016), may coincide with appreciation of the overall reading experience. It may lead readers to see that the text is *well made* and appreciate its craftsmanship or "poetic" aspects in the ancient sense of the term (cf. "poietical" knowledge, referring to how something is made, Atkins, 1934).

Rereading benefits comprehension

To conceptualize the emergent effect of rereading we propose to consider a wider perspective, looking beyond literary theory. As mentioned before, rereading is studied in various disciplines. First, research in education studies suggest that what *is* emerging on rereading could be meta-comprehension accuracy (Rawson, Dunlosky, & Thiede, 2000), that is, readers' own judgment of how well they comprehended a text. As Rawson et al. suggested, the fact that previous research has found comprehension of a text increases upon rereading points to the possibility that texts are not processed in their entirety on first reading. For instance, Millis, Simon, and Tenbroek (1998) show that text base construction dominates the first reading and constructing situation models dominates the second reading. Because fewer disruptions occur in text base processing upon second reading, readers can relocate their cognitive resources, enabling them to construct a more complete situation model upon second reading. As a result they will understand the text better or more fully and acquire higher levels of meta-comprehension accuracy. It might be that improved understanding facilitates appreciation, which would then explain the emergent appreciation of rereading for any text, including literature.

Converging evidence for this role of the reallocation of cognitive resources when rereading can be found in studies that show that increases in *processing fluency* (i.e., ease of processing) are related to enhanced appreciation (Reber, Schwarz, & Winkielman, 2004). Evidence suggests that this "hedonic fluency model" holds for perceiving visual art (Kuchinke, Trapp, Jacobs, & Leder, 2009; Winkielman & Cacioppo, 2001): The easier paintings are processed, the higher viewers' positive aesthetic emotions and appreciation. Comparable findings have revealed the aesthetic attraction of prototypicality of images (images "already seen"), colors, and paintings (Farkas, 2002; Martindale & Moore, 1988) and increases in liking and canonization due to the mere exposure effect (cf. Cutting, 2006; Zajonc, 1968). These studies in aesthetic perception might lead us to speculate that similar processes are at work in responses to literature: The easier it is to process a text (e.g., on second reading), the more readers will appreciate it.

In sum, it could be (1) readers' estimation of the qualities of literary style or (2) a more common increase in comprehension that causes increased appreciation. The two propositions are not *necessarily* mutually exclusive: It may be that increased appreciation on rereading is caused by literariness *and* increased comprehension, with readers' more comprehensive understanding accounting for aspects of the literary style only on the second reading. Whether these processes occur may be determined by many factors, including individual differences. Predictions regarding readers' appreciation and understanding of literariness may need to include their competence and experience.

Reading experience

Dixon et al. (1993) found that results for the frequent readers in their sample revealed an emergent effect; those of the infrequent readers did not. Even so, we assume that reading experience can intervene in two opposite ways. Highly experienced readers may appreciate foregrounding aspects in a text more because they recognize the significance of these features more readily than readers with less experience. That would mean they benefit more from rereading (as they seem to have done in Dixon et al.'s study). On the other hand, one could easily reason the other way around: Experienced readers would not need a rereading because they have already come to value the significance of foregrounded text features on their first reading. The same may hold, we believe, for comprehension: Experienced readers' comprehension would gain little from rereading because it was already high on first reading. In any event it seems likely that these relations are not linear and that they depend on the level of challenge that the text presents to its readers. It may be that they reveal a reversed U-curve, with both low and high levels of reading experience leading to low appreciation and with intermediate levels leading to maximal appreciation (cf. Berlyne, 1970).

In conclusion, it may be that a text's level of foregrounding determines whether rereading leads to increased appreciation; this would support the literariness hypothesis. Second, it may also be argued that it is increased comprehension that leads to increases in appreciation, irrespective of the level of foregrounding. This would support the hypothesis that the emergent effect is a much more common phenomenon. A third possibility is that these two factors work together and are both associated with higher levels of appreciation on second reading, as compared with evaluations on the first. Finally, we assume that reader experience counts, one way or the other. For now, we base our predictions on the findings of Dixon et al. (1993) and assume that experienced readers may be more appreciative of foregrounding in the text than less experienced readers and hence the emergent effect will be larger for them.

Present research

We conducted three experiments investigating the relation between the level of literariness and increases in appreciation, using linear mixed models and Dixon et al.'s (1993) rereading paradigm, with participants reading either an original literary text or a manipulated version with reduced levels of literariness. We made some adjustment on the work of Dixon et al. First, we varied the type of text

manipulation. In Experiment 1 the rewriting strictly followed Dixon et al. (1993), albeit applied to the Dutch translation of the Borges story rather than an English one. In Experiments 2 and 3, however, we introduced a replicable, pervasive, and quantifiable approach to literariness by conceptualizing it as foregrounding. We based the manipulation on foregrounding text analyses, using Miall and Kuiken's (1994) systematic method, allowing a sentence-by-sentence weighing of the level of foregrounding and thus a sentence-by-sentence rewriting. In principle this approach can be used in any literary text and enables testing whether the levels of foregrounding in two versions of the same text are significantly different.

Furthermore, in addition to previous research, we examined whether increases in appreciation could be related to increases in comprehension. To this end we included a set of items measuring perceived comprehension (Kuijpers, 2014) as a time-varying covariate in our linear mixed model analyses. The items we used are, of course, a *subjective* measure of comprehension. Criteria for an objective comprehension measure for literary texts are, at best, debatable because literary conventions would have it that one text can have different legitimate interpretations for different readers (cf. the polyvalence of literature, e.g., Schmidt, 1982). Moreover, chances of test (learning) effects on a repeated comprehension measure are considerable. Instead, we used a subjective measure, assuming it is a readers' *feeling* that they have a better understanding of a story that might affect appreciation, and not necessarily their actual and accurate understanding. In this, we follow previous research relating rereading to meta-comprehension accuracy (i.e., people's own judgment of how well they comprehended a text; cf. Rawson et al., 2000). We wanted to know whether an increase in perceived comprehension explains an increase in appreciation, dependent or independent of the level of literariness. For that purpose we entered perceived comprehension into the model first and then added "condition" (high literariness vs. low literariness), checking whether the level of literariness can explain additional variance in possible increases in appreciation. To assess whether there is conceptual overlap between the two constructs of appreciation and comprehension that could bias our results, we conducted a factor analysis (i.e., principal components analysis with oblique rotation based on eigenvalues greater than one, suppressing factor loadings lower than .45) on appreciation and comprehension items combined. In all three studies (and for both first and second readings) we found that the two concepts reliably fall into two separate factors. There was only one case in which a comprehension item ("The story was an easy read") loaded on the appreciation factor, namely in Experiment 3 on the first reading. Based on the fact that only one item in only one of six factor analyses double loaded, we believed it safe to assume that the two concepts of appreciation and comprehension are sufficiently independent of one another.

To minimize the effects of social desirability, we replaced the self-report measures for reading frequency as used in Dixon et al. (1993), with the Author Recognition Test (ART; Koopman, 2015; Mol & Bus, 2011; Stanovich & West, 1989). Self-reported reading frequency could be exaggerated by participants. The ART discourages such responses. It involves a list of author names, both real names (or pseudonyms) and fake names. Participants are instructed to indicate the author names they recognize but are also told that there are false names on the list. The number of real authors they indicate determines their overall print exposure score. When a reader marks more than three false names they are excluded from the analyses.

We examined the role of reader experience in a different way than Dixon et al. In their study the sample was divided based on self-reported reading frequency using a median split. We uses the same procedure here and report confidence intervals and sensitivity index to allow for comparison of patterns with Dixon et al.'s study. For this, we subtracted the depth of appreciation score for the manipulated story from that of the original story. We did this per group of readers and then subtracted one group from the other. However, as the procedure of using a median-split approach results in loss of nuance and statistical power, we also ran a series of linear mixed model analyses in which we considered print exposure as a covariate. This procedure allowed us to investigate the influence of perceived comprehension on increased appreciation after rereading while taking both reader experience and condition (level of literariness) into account.

As a final adjustment on previous work, we investigated the role of test reactivity by having half of the participants reread the text 1 week later (as in Mills et al. 1998) and the other directly after the first evaluation (as in Dixon et al., 1993). This resulted in a 2 (original story vs. manipulated story) × 2 (direct rereading vs. delayed rereading) design for Experiment 1.

Overview

We hypothesize that texts with narratorial ambiguity (Experiment 1) and with higher foregrounding scores (i.e., higher levels of literariness; Experiments 2 and 3) will elicit a greater increase in appreciation (Hypothesis 1). Second, we expect increases in appreciation to be related to increases in perceived comprehension (Hypothesis 2). Finally, we hypothesized that readers' print exposure will moderate the effects of text on appreciation: the higher print exposure, the higher the increase in appreciation. (Hypothesis 3). The first study replicates and extends the experiment by Dixon et al. (1993). We manipulated a Dutch translation of Borges' *Emma Zunz* (2003) in the same way Dixon et al. originally did. In the second study we investigated the effects of manipulating a fragment of a highly foregrounded story (Salman Rushdie's *Midnight's Children*) on perceived comprehension and appreciation using the rereading paradigm. In the third study we attempted to replicate the results of the second study with a different sample and a different story (Ambrose Bierce's *An occurrence at Owl Creek Bridge*).

Experiment 1

Methods

Participants

Ninety-seven students of the Languages and Literature Department of Utrecht University participated in this experiment voluntarily. Seventy-nine were female and 16 male. Ages ranged from 17 to 39 ($M = 19.09$, $SD = 2.97$). Scores on the ART can vary between 0 and 30. In our sample scores ranged between 2 and 23, with a mean of 9.78 ($SD = 4.86$).

Materials

We used the same story as Dixon et al. (1993), Borges' *Emma Zunz*, be it in Dutch translation, and applied the same text manipulation, that is, we removed all explicit references to the narrator's lack of knowledge or uncertainty. In doing so we modified the same 17 passages that Dixon et al. changed. In Table 1 we provide two examples from Dixon et al.'s study and the Dutch translations and manipulation, we used in the present study. The original story counted 2,130 words and the manipulated version had 2,078.

Table 1. Example of Manipulation of Jorge Luis Borges' *Emma Zunz*.

Original text (Dixon et al., 1993)	Modified version (Dixon et al., 1993)
Perhaps it comforted her to verify in the insipid movement along the streets that what had happened had not contaminated things	It comforted her to verify in the insipid movement along the streets that what had happened had not contaminated things. Only once did Emma Zunz think about the dead man who motivated the sacrifice.
Did Emma Zunz think once about the dead man who motivated the sacrifice?	
Original text (Dutch translation)	**Modified version (Dutch translation)**
Misschien bemoedigde het haar te merken dat het gebeurde de dingen, in de suffe drukte op straat, niet had aangetast.	Het bemoedigde haar te merken dat het gebeurde de dingen, in de suffe drukte op straat, niet had aangetast.
Dacht Emma Zunz in die tijd [...] *één enkele maal* aan de dode die aanleiding was tot haar offer?	In die tijd [...] dacht Emma Zunz slechts *één enkele maal* aan de dode die de aanleiding was tot haar offer [...]

Procedure

The experiment combined a 2 (original story versus manipulated story) × 2 factor (direct rereading vs. rereading after a week) between-subjects design with a within-subjects design (ratings after first reading vs. ratings after second reading). Participants were randomly assigned to one of four conditions. After reading the story they were asked to fill out a questionnaire. Two sets of items were used to measure respectively perceived comprehension (first reading, α = .78; second reading, α = .79) and appreciation (first reading, α = .87; second reading, α = .89). We conducted additional reliability analyses on the change scores of the seven appreciation items to assess whether they provide an internally consistent index of *depth of appreciation*. They do (α = .74). There were five items included on the perceived comprehension measure, namely "I could follow the leading thread of the story," "I found the story was understandable," "The story was an easy read," "I did not stumble upon difficult words or sentences," and "I thought the style of writing was accessible." Seven items were included in the appreciation measure, namely "I thought it was fun to read this story," "I thought the story was written well," "I read the story with great interest," "I thought the story was beautiful," "I thought it was an exciting story," "I thought it was an emotional story," and "I thought it was rewarding to read this story." It should be noted that Dixon et al. used three different items to measure appreciation, pertaining to perceived quality, reported enjoyment, and readiness to recommend the text to a friend. In our measure we added items capturing two aspects of participants' notion of appreciation: perceived quality and pleasure of reading. We used this set of seven items as their reliability was demonstrated in previous studies (Kuijpers, 2014).

We treated perceived comprehension and appreciation as hypothetically related but conceptually distinct: One can understand a poem, story, or painting without appreciating it. Similarly, it is possible to find a literary text or visual artwork appealing even in the absence of (full) understanding (cf. Leder & Nadal, 2014, p. 447). The results of the factor analyses confirm that we are dealing with two conceptually distinguishable constructs.

After filling out the questionnaire, participants were asked to read the story again, either instantly or after receiving an e-mail a week later. When they were done rereading, they were asked to fill out the same questionnaire as before. In addition, they completed a Dutch version of the ART (Koopman, 2015) and provided sociodemographic information about themselves (i.e., their age and gender). All participants, in this study and the following, were debriefed about the purposes and the results of the research.

Statistical analyses

We conducted a series of mixed model analyses in all three studies with appreciation as the outcome variable. Such analyses allowed us to test a repeated measures mixed design (i.e., rereading by condition by print exposure) and include a time-varying covariate (i.e., perceived comprehension; West, Welch, & Galecki, 2015). As fixed effects we entered condition and print exposure and our dependent variable of perceived comprehension into our models. As random effects we included intercepts for subjects. The first mixed model analysis included perceived comprehension as repeated measures dependent variable. In the second model we added print exposure as independent variable to explore whether it could explain any additional variance. In the third model we added condition as independent variable to test for additional variance. The significance threshold in all analyses was set at .05.

Results

Increases in appreciation

To test the first hypothesis that the original text elicits greater increases in appreciation than the manipulated text, we conducted a repeated-measures General Linear Model (GLM). Table 2 shows the means and standard deviations of appreciation and perceived comprehension per condition.

Table 2. Means and Standard Deviations of Scores on Comprehension and Appreciation per Condition and per Reading of Experiment 1.

	Original Story Direct Rereading (High Literariness)						Manipulated Story Direct Rereading (Low Literariness)					
	First Reading		Second Reading		Difference Score		First Reading		Second Reading		Difference Score	
	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI
Comprehension	4.27	3.85,	4.76	4.37,	.49	.18,	3.89	3.38,	4.81	4.39,	.92	.56,
	(1.17)	4.69	(1.07)	5.14	(.84)	.79	(1.33)	4.39	(1.09)	5.22	(.94)	1.28
Appreciation	3.79	3.37,	3.98	3.59,	.19	−.06,	3.76	3.31,	4.17	3.70,	.41	.16,
	(1.17)	4.21	(1.07)	4.37	(.71)	.45	(1.18)	4.21	(1.23)	4.64	(.66)	.67
	Original Story Delayed Rereading (High Literariness)						Manipulated Story Delayed Rereading (Low Literariness)					
	First Reading		Second Reading		Difference Score		First Reading		Second Reading		Difference Score	
	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI
Comprehension	4.43	3.82,	4.55	3.86,	.13	−.35,	4.29	3.82,	4.18	3.72,	−.11	−.59,
	(1.14)	5.03	(1.29)	5.24	(.89)	.60	(.89)	4.76	(.86)	4.64	(.90)	.37
Appreciation	3.91	3.54,	3.76	3.16,	−.15	−.50,	3.76	3.24,	3.64	3.02,	−.13	−.66,
	(.70)	4.28	(1.13)	4.36	(.66)	.19	(.97)	4.28	(1.17)	4.26	(1.01)	.41

Notes. These are the actual means, not adjusted for any covariates.

The analysis revealed no significant difference between appreciation scores on first and second reading ($F(1,90) = .93$, $p = .34$, $\eta_p^2 = .01$). This holds across conditions. In other words we can reject the hypothesis that the original story elicited greater increases in appreciation. Also, there was no difference in appreciation between the group who immediately reread the story and the group who reread the story 1 week later. Based on these results we can assume that retest effects do not occur in the immediate rereading conditions.

Role of comprehension

To test our second and third hypotheses we conducted a series of mixed model analyses. The results of the analyses can be found in Table 3. The first mixed model analysis with appreciation as a time-varying dependent variable and perceived comprehension as a time-varying covariate revealed a significant positive relation between perceived comprehension and appreciation: As perceived comprehension increases from first to second reading, so does appreciation ($R^2 = .22$). When adding

Table 3. Fixed Effect Estimates and Variance-Covariance Estimates for Repeated Measures Mixed Models Experiment 1.

	Estimator	Standard Error	t(df)
Model 1			
Fixed effect			
Comprehension	.441***	.058	7.65 (179)
Random effect			
Intercept (subjects)	.690	.119	
Model 2			
Fixed effects			
Comprehension	.473***	.068	6.98 (163)
Print exposure	−.066**	.019	−3.32 (121)
Random effect			
Intercept (subjects)	.633	.111	
Model 3			
Fixed effects			
Comprehension	.461***	.072	6.43 (161)
Print exposure	−.064**	.021	−3.04 (117)
Condition	−002	.291	−.007 (119)
Random effect			
Intercept (subjects)	.656	.117	

***$p < .00$, **$p < .01$

the covariate of print exposure to the second model, perceived comprehension was still significantly related to increases in appreciation, and we see that print exposure explains some additional variance: As print exposure decreases, appreciation increases, albeit very little ($R^2 = .03$). The third model, in which we added condition as independent variable, shows that the significant association between perceived comprehension and appreciation is still there, as is the effect of print exposure. However, condition does not explain any additional variance. Hence, we can accept the hypothesis that increases in appreciation are more related to increases in perceived comprehension than to the level of literariness.

Role of print exposure

The mixed models show that as print exposure decreases, appreciation increases. However, to be able to compare our findings with those of Dixon et al. (1993), we also conducted a median split on our sample, resulting in two groups: one with low print exposure (infrequent) readers versus high print exposure (frequent) readers. Then we compared their "depth of appreciation" scores. It must be noted that we need to be careful in our comparisons because we used another measure for reading frequency. The results of these analyses can be seen in Figure 1 below. Dixon et al. found that frequent readers showed a substantial depth of appreciation for rereading the original story but not for rereading the manipulated story, whereas infrequent readers were unaffected by the manipulation. In contrast, our results suggest a substantial depth of appreciation among the infrequent readers of the original story and even more for the manipulated one, whereas the experienced readers were unaffected by the manipulation. To facilitate the comparison we also calculated a sensitivity index as Dixon et al. did, we subtracted the depth of appreciation measure for the manipulated story from that for the original story (Dixon et al., 1993, p. 28). This index gives an estimation of participants' sensitivity to the manipulation. The greater the value, the greater the sensitivity. A negative value indicates that the depth of appreciation for the manipulated story was smaller than the depth of appreciation for the original story or that there was a decrease in
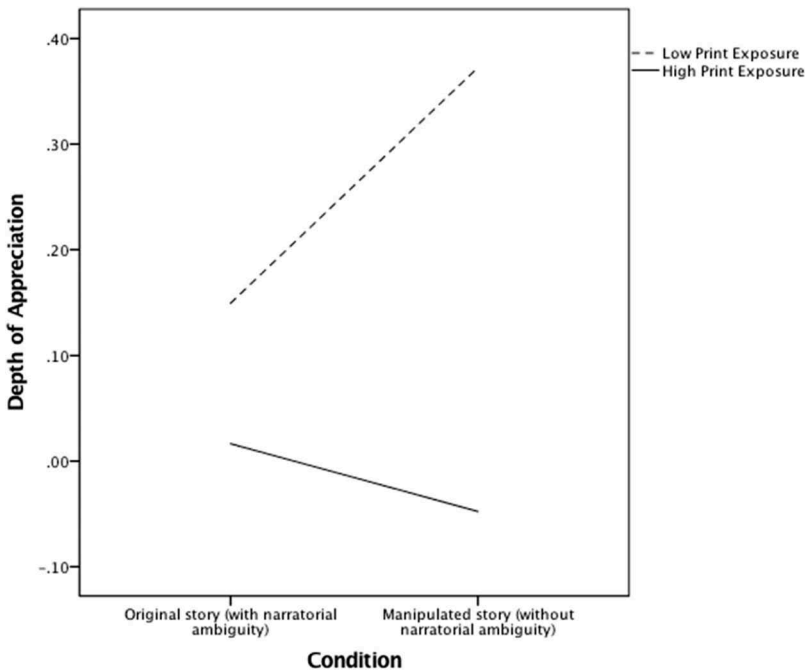

Figure 1. Depth of appreciation by low print exposure and high print exposure readers with and without narratorial ambiguity.

appreciation from one reading to the next in one of the conditions, and thus the depth of appreciation scores would already be negative.

The sensitivity index for the inexperienced readers in this study was –.22, whereas the one for the experienced readers was –.05. Both numbers are relatively small, especially compared with Dixon et al.'s results (i.e., 2.78 for frequent readers and –.36 for infrequent readers). The difference between experienced and inexperienced readers is almost negligible (–.19), suggesting no real difference between the two groups of readers, and hence reading experience did not seem to be a requirement to enable the appreciation of narratorial ambiguity. Moreover, an increase in appreciation from first to second reading did occur for the *inexperienced* readers in both conditions—contrary to Dixon et al.'s results—and for the experienced readers only when reading the original story, which tells us that something apart from the narratorial ambiguity increased appreciation. Based on the results of the mixed model analyses, we suggest an increase in comprehension occurs from first to second reading.

Discussion

The results from Experiment 1 seem to suggest that it was not the presence of a literary device in the text that led to increases in appreciation from first to second reading. Instead, it may be suggested that an increase in perceived comprehension between readings leads to increases in appreciation, irrespective of the level of literariness. As to the role of print exposure, low print exposure participants experience greater increases in appreciation than high print exposure participants. It may be that we did not find any effect of the text manipulation because it was not strong enough. To examine this possibility, we ran a second study in which we use the rereading paradigm on another literary text and a different type of manipulation. We still wanted to focus on the concept of literariness but also wished to create two versions of one text that differed *significantly* from one another, more so than the two versions of *Emma Zunz* in Experiment 1. As explained in the Introduction, we conceptualize literariness as foregrounding, which can, potentially, be more pervasively present throughout a text than narrational ambiguity. We used a text that matches that criterion well, an original text by Rushdie, and removed as much foregrounding as possible. Thus, we created a set of stimuli that would allow us to test our conclusions from Experiment 1. We used the same measures and design. However, we dropped the delayed rereading condition since our results showed no differences between delayed and immediate rereading.

Experiment 2

Methods

Participants

Fifty-one book club members participated in this experiment. Their ages ranged from 21 to 71 with a mean age of 49.35 and a standard deviation of 16.95. Forty-one participants were women and 10 men. The participants were recruited via an Internet search (i.e., Facebook, LinkedIn, Goodreads) for interested book clubs and via public libraries. The reason for soliciting participants in this fashion was to increase the ecological validity of the experiment, assuming that this particular sample of participants was used to rereading and reflecting on their reading and interested in more challenging texts (making them comparable with students of literature). The range of print exposure scores within this sample was 20 with a minimum of 4 and a maximum of 24, the mean score was 16.27, and the standard deviation was 4.46.

Materials

We applied our manipulation throughout the text. We used a literary text that is highly foregrounded: the first four pages of Salman Rushdie's novel *Midnight's Children*. The selected text

Table 4. Example of Manipulation of Salman Rushdie's *Midnight's Children*.

Original (high foregrounding condition)

I was born in the city of Bombay…once upon a time. No, that won't do, there's no getting away from the date: I was born in Doctor Narlikar's Nursing Home on August 15th, 1947. And the time? The time matters, too. Well then: At night. No, it's important to be more… On the stroke of midnight, as a matter of fact. Clockhands joined palms in respectful greeting when I came. Oh, spell it out; spell it out: At the precise instant of India's arrival at independence, I tumbled forth in the world. There were sobs. And, outside the window, fireworks and crowds. Thanks to the occult tyrannies of those blandly saluting clocks I had been mysteriously handcuffed to history, my destinies indissolubly chained to those of my country.

Manipulated version (low foregrounding condition)

I was born on 15 August 1947 in Bombay in Doctor Narlikar's Nursing Home. At midnight, at precisely the same time that India became independent, I was born. That night there were fireworks and people went out on the streets in crowds, but there were also people crying. When I think about the night of my birth, I see that my fate is tied to the fate of my country, because I was born at that precise moment when India became independent.

can be read as a story on its own, and it is a text that contains many foregrounded elements (Rushdie, 2003) . The expectation was that this would lead to two radically different versions of the story after manipulation. In Table 4 an example can be found of how the rewritten version compares with the original. The low foregrounding story counts 1,275 words and the high foregrounding story 1,337 words.

As can be seen in the example above, we left out deviating use of punctuation, neutralized the narratorial ambiguity, left out the direct questions to the reader, converted metaphoric expressions into neutral language, and made sure that no repetition occurred within sentences or across sentences (i.e., no alliteration, or repetition of words or phrases). We made sure we introduced alterations at all the three textual levels investigated in Miall and Kuiken's approach, that is, at the phonetic, grammatical, and semantic level (cf. Miall & Kuiken, 1994). We tried to neutralize foregrounding as far as possible without actually changing the content of the story. Consequently, there is still some foregrounding left in the manipulated version. However, for the purpose of this study it sufficed to have two versions of the story that differed to a significantly large extent in their levels of foregrounding.

Foregrounding text analyses of the original and the manipulated version of the Rushdie story were conducted using Miall and Kuiken's method (Miall & Kuiken, 1994). The stories were divided into segments, mostly following sentence length or phrase length. Two raters (the first author and a published creative writer with a master's degree in literary studies) worked independently, analyzing the phonetic, grammatical, and semantic text levels of these segments and counted the number of foregrounded features per category per segment. The coding scheme was agreed on by the two raters before analyzing the texts. The interjudge agreement was calculated by correlating the frequencies of all the features combined per segment by each judge. The mean correlation was highly significant for both versions of the story ($p < .001$): for Rushdie's original version of *Midnight's' Children, r* (55) = .94; for our manipulated version of *Midnight's Children, r*(61) = .66. The sum of each category of the original story was compared with the sum of that same category in the manipulated story. Results of the foregrounding text analyses and univariate analyses can be found in Table 5.

We chose not to control for syllable length per sentence when calculating the foregrounding scores because sentence length was also manipulated. To improve readability, some longer sentences

Table 5. Mean Number of Syllables per Segment per Story (Original vs. Manipulated), Sum Foregrounding Scores per Category (Phonetic, Grammatical, or Semantic) per Story in Experiment 1.

	Average Number of Syllables per Segment	Sum Phonetic Scores	Sum Grammatical Scores	Sum Semantic Scores
Rushdie Original	40.49	178*	157***	153***
Rushdie Manipulated	31.28	143	54	51

***$p < .001$, *$p < .05$

were divided into several shorter sentences, which according to Van Peer (1986) also reduces literariness.

The results of the analyses show that we were successful in our manipulation on all three text levels: phonetic $F(1,114) = 6.88$, $p = .010$, $\eta_p^2 = .057$; grammatical $F(1,114) = 17.73$, $p < .001$, $\eta_p^2 = .135$; and semantic $F(1,114) = 18.36$, $p < .001$, $\eta_p^2 = .139$. From this, we conclude that our manipulation succeeded in creating two versions of Rushdie's story differing significantly in terms of foregrounding.

Procedure

The experiment combined a two-factor between-subjects design (condition: low foregrounding vs. high foregrounding) with a within-subjects design (ratings after first reading vs. ratings after second reading). Participants were randomly assigned to one of the two texts. After reading the story they were asked to fill out a questionnaire. Two sets of items were used to measure, respectively, perceived comprehension (first reading, $\alpha = .85$; second reading, $\alpha = .89$), and appreciation (first reading, $\alpha = .88$; second reading, $\alpha = .89$). The Cronbach's α of the change score (i.e., the difference score between first and second reading appreciation) was .82.

Results

Increases in appreciation

To test the first hypothesis that the original text elicits greater increases in appreciation than the manipulated text, we conducted a repeated measures GLM. Table 6 shows the means and standard deviations of appreciation and perceived comprehension per condition (low foregrounding vs. high foregrounding) and per reading (first vs. second).

The analysis revealed a significant difference between appreciation scores on first and second reading, $F(1,48) = 10.28$, $p < .01$, $\eta_p^2 = .18$, independent of the text that was read. Participants appreciated *both* text versions significantly more after the second reading compared with the first reading. The hypothesis that only high levels of literariness lead to greater increases in appreciation needs to be rejected.

Role of comprehension

To test our second and third hypothesis we conducted a series of mixed model analyses (Table 7). The first mixed model analysis with appreciation as a time-varying dependent variable and perceived comprehension as a time-varying covariate revealed a significant positive relation between perceived comprehension and appreciation: As perceived comprehension increases, so does appreciation ($R^2 = .38$). When adding the covariate of print exposure to the second model, perceived comprehension is still significantly related to appreciation, but print exposure did not explain any additional variance. The third model, in which we added condition as independent variable, shows that the

Table 6. Means and Standard Deviations of Scores on Comprehension and Appreciation per Condition and per Reading for Experiment 2.

	Manipulated Story (Low Foregrounding)						Original Story (High Foregrounding)					
	First Reading		Second Reading		Difference Score		First Reading		Second Reading		Difference Score	
	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI
Comprehension	5.33 (1.11)	4.86, 5.81	5.60 (.94)	5.20, 5.99	.27 (.85)	−.09, .62	3.78 (1.37)	3.23, 4.34	4.56 (1.45)	3.97, 5.15	.78 (.89)	.42, 1.13
Appreciation	4.15 (1.26)	3.62, 4.69	4.57 (1.14)	4.08, 5.05	.41 (.97)	.01, .82	4.19 (1.30)	3.66, 4.71	4.60 (1.32)	4.07, 5.13	.41 (.85)	.07, .75

Notes. These are the actual means, not adjusted for any covariates.

Table 7. Fixed Effect Estimates and Variance-Covariance Estimates for Repeated Measures Mixed Models Experiment 1.

	Estimator	Standard Error	t(df)
Model 1			
Fixed effect			
Comprehension	.530***	.078	6.81 (98)
Random effect			
Intercept (subjects)	.887	.207	
Model 2			
Fixed effects			
Comprehension	.585***	.093	6.32 (92)
Print exposure	−.051	.035	−1.48 (62)
Random effect			
Intercept (subjects)	.881	.206	
Model 3			
Fixed effects			
Comprehension	.669***	.095	7.00 (88)
Print exposure	−.057	.033	−1.75 (63)
Condition	−.706*	.299	−2.36 (67)
Random effect			
Intercept (subjects)	.758	.183	

***$p < .00$; *$p < .05$

significant association between perceived comprehension and appreciation is still there and that condition seems to explain only a little additional variance.

Role of print exposure
The mixed models show that print exposure did not have an effect on increases in appreciation within the model. However, to compare the two groups of readers—high print exposure (frequent) readers versus low print exposure (infrequent) readers—we conducted a median split on our sample
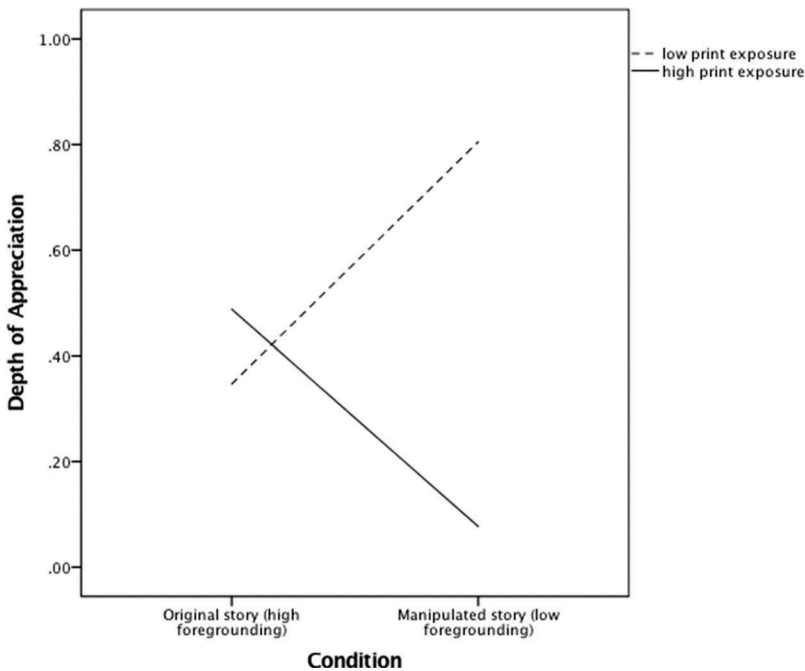

Figure 2. Depth of appreciation by low print exposure and high print exposure readers with and without foregrounding.

and investigated the interaction between rereading and reading experience separately. The results are presented in Figure 2.

In this study we see a pattern similar to what Dixon et al. (1993) found: Responses of experienced readers show a substantial depth of appreciation for the original story but not for the manipulated version (sensitivity index, .41). However, for the inexperienced readers we also find a substantial depth of appreciation for both conditions (sensitivity index, –.46). The difference in terms of sensitivity for foregrounding between the two groups is .87, which is relatively small again in comparison with Dixon et al.'s results. Together with the results from our first study, it seems we are able to replicate the results of Dixon et al., but only for the experienced readers. We see that inexperienced readers actually report a greater increase in appreciation when reading the low foregrounding version.

Discussion

This study investigated the relationship between perceived comprehension and appreciation upon rereading of two texts differing in level of literariness, operationalized as foregrounding. The results show that the increase in appreciation upon second reading occurs in both conditions, independent of the level of literariness. Again, we need to reject the hypothesis that only the highly literary text would elicit higher levels of appreciation upon rereading. The results of the mixed models suggest rather that perceived comprehension is the best predictor of increased appreciation. Therefore, the hypothesis that increases in appreciation upon rereading are more related to increases in perceived comprehension than to the presence of literariness in a text can be accepted. However, the hypothesis that high print exposure is associated with greater increases in appreciation from first to second reading needs to be rejected in this study, since we found that print exposure did not have any effect on appreciation. One possible reason why we did not find an effect of print exposure in the linear mixed model but did find such an effect in Experiment 1 is that a sample of book club members was used, which resulted in a slightly smaller degree of variation on the ART.

Therefore, we decided to conduct Experiment 3 with a sample we expected to vary more on their print exposure scores, hoping this would make it more likely we would detect the role of this variable, if any. In this third study we also correlated readers' previous print exposure as measured by the ART with a simple measure of reading frequency to check whether the failure to replicate the results of Dixon et al. (1993) was because we used a different measure. In Experiment 3 the same experimental design and manipulation are used as in Experiment 2 but with different stimulus material and a different sample. In the two previous studies we used texts with high (Borges) to very high (Rushdie) levels of foregrounding. Our aim in Experiment 3 was to investigate whether our results would also hold for a text with relatively low levels of foregrounding, but still of noticeable literary value.

Experiment 3

Methods

Participants

Students enrolled in a course on audience reception research at Utrecht University were asked to recruit participants for this study among family and friends. This resulted in a sample of 49 participants. Ages ranged from 17 to 78 ($M = 35$, $SD = 16.67$). Fifteen participants were male and 33 female (1 case was missing). ART scores ranged from 0 to 22, with a mean of 10.12 ($SD = 4.78$). This range is comparable with the one in Experiment 2, and therefore our expectation of more variation on the ART through broader sampling was not met. This is an unfortunate but not a fundamental problem, however, since it will make a comparison of the groups across studies more valid.

Table 8. Examples of Manipulations per Version of Ambrose Bierce's *An Occurrence at Owl Creek Bridge*.

Original (high foregrounding condition)

A man stood upon a railroad bridge, looking down into the swift water twenty feet below. The man's hands were behind his back, the wrists bound with a cord. A rope closely encircled his neck. It was attached to a stout cross-timber above his head and the slack fell to the level of his knees. Some loose boards laid upon the sleepers supporting the metals of the railway supplied a footing for him and his executioners–two private soldiers of the Federal army, directed by a sergeant who in civil life may have been a deputy sheriff. At a short remove upon the same temporary platform was an officer in the uniform of his rank, armed. He was a captain. A sentinel at each end of the bridge stood with his rifle in the position known as "support," that is to say, vertical in front of the left shoulder, the hammer resting on the forearm thrown straight across the chest–a formal and unnatural position, enforcing an erect carriage of the body. It did not appear to be the duty of these two men to know what was occurring at the center of the bridge; they merely blockaded the two ends of the foot planking that traversed it.

Manipulated version (low foregrounding condition)

A man stood upon a railroad bridge, looking down on a river. His wrists were bound with a cord behind his back. A rope was tied closely around his neck. I was attached to a beam above his head. He was standing on some loose boards that were laid out over the railway. There were two soldiers and a sergeant standing next to him. At a little distance on the same platform, there stood an armed captain. On both ends of the bridge there were two soldiers standing guard, with their rifles in the 'support' position. This meant that they let their rifles rest on their lower arms, with the barrel leaning against their left shoulders. This is a position that forces a person to stand straight. It was their duty to block the entrances to the bridge, but they could not see what was happening on the bridge.

Materials

For the purpose of this study a short story, *An Occurrence at Owl Creek Bridge* by Ambrose Bierce, was manipulated (Bierce, 1975). The story was chosen because of the results of a pilot study with 10 expert readers who had to evaluate six different stories on their literariness and whether or not they found the stories interesting. Participants scored the Bierce story highest on interestingness and considered it to be literary. The original story was used in the high foregrounding condition and was rewritten to remove as much foregrounding as possible for the low foregrounding condition. Table 8 shows an example of the manipulations in both stories. The low foregrounding story counts 771 words and the high foregrounding story 1,064 words.

As can be seen in Table 8 we left out narratorial commentary, left out adjectives, converted metaphors into less deviating language, and made sure that no repetition occurred (cf. Miall & Kuiken, 1994). As in Experiment 2 there is still some foregrounding left in the manipulated version. Again, we carried out the neutralization of foregrounding as far as possible without changing the content of the story.

Foregrounding text analyses of the original and the manipulated version of the Bierce story were conducted using Miall and Kuiken's method (Miall & Kuiken, 1994). The interjudge agreement was calculated again by correlating the frequencies of foregrounding features per segment by each judge (i.e., the same as in Experiment 2). The mean correlations were highly significant for both versions of the story (p < .001): Bierce's original version of *An occurrence at Owl Creek Bridge*, $r(54) = .76$, and our manipulated version of *An occurrence at Owl Creek Bridge*, $r(55) = .60$. Results of the text analyses, univariate analyses, and the average number of syllables per segment can be found in Table 9. The results of the analyses show that we were successful in our manipulation on all three text levels: phonetic $F(1,107) = 31.57$, $p < .001$, $\eta_p^2 = .228$; grammatical $F(1,107) = 19.35$, $p < .001$, $\eta_p^2 = .153$; and semantic $F(1,107) = 21.05$, $p < .001$, $\eta_p^2 = .164$.

Table 9. Mean Number of Syllables per Segment per Story (Original vs. Manipulated), Sum Foregrounding Scores per Category (Phonetic, Grammatical, or Semantic) per Story in Experiment 2.

	Average Number of Syllables per Segment	Sum Phonetic Scores	Sum Grammatical Scores	Sum Semantic Scores
Bierce original	31.22	232***	77***	61***
Bierce manipulated	21.27	125	26	16

***p < .001

Procedure

We followed the same procedures as was used in Experiment 2. For both perceived comprehension (first reading, α = .71; second reading, α = .73) and appreciation (first reading, α = .89; second reading, α = .90) we found reliable ratings. The Cronbach's α of the change score between first and second reading appreciation was .71.

Results

Increases in appreciation

To test the first hypothesis that the original text elicits greater increases in appreciation than the manipulated text, we conducted a repeated measures GLM (Table 10). The analysis revealed no significant difference between appreciation scores on first and second reading, $F(1,47)$ = .009, p = .93, η_p^2 = .00. This holds for both conditions.

Role of comprehension

To test our second and third hypothesis we conducted a series of mixed model analyses (Table 11). The first mixed model analysis with appreciation as a time-varying dependent variable and perceived comprehension as a time-varying covariate revealed a significant positive relation between perceived comprehension and appreciation: As perceived comprehension increases, so does appreciation (R^2 = .18). When adding the covariate of print exposure to the second model, perceived

Table 10. Means and Standard Deviations of Scores on Comprehension and Appreciation per Condition and per Reading for Experiment 3.

	Manipulated Story (Low Foregrounding)						Original Story (High Foregrounding)					
	First reading		Second Reading		Difference Score		First Reading		Second Reading		Difference Score	
	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI
Comprehension	4.23	3.78,	4.68	4.25,	.45	.13,	3.85	3.41,	4.42	4.02,	.56	.29,
	(1.12)	4.68	(1.06)	5.10	(.79)	.76	(1.04)	4.30	(.92)	4.82	(.64)	.84
Appreciation	3.77	3.28,	3.69	3.21,	−.08	−.40,	3.34	2.85,	3.40	2.91,	.06	−.25,
	(1.22)	4.26	(1.18)	4.16	(.79)	.24	(1.11)	3.82	(1.12)	3.88	(.72)	.37

Notes. These are the actual means, not adjusted for any covariates.

Table 11. Fixed Effect Estimates and Variance-Covariance Estimates for Repeated Measures Mixed Models with Print Exposure Experiment 2.

	Estimator	Standard Error	t(df)
Model 1			
Fixed effect			
Comprehension	.374**	.105	3.57 (94)
Random effect			
Intercept (subjects)	1.026	.236	
Model 2			
Fixed effects			
Comprehension	.421**	.124	3.41 (86)
Print exposure	−.042	.034	−1.22 (57)
Random effect			
Intercept (subjects)	1.010	.237	
Model 3			
Fixed effects			
Comprehension	.414**	.125	3.31 (84)
Print exposure	−.040	.034	−1.17 (55)
Condition	.142	.326	.44 (56)
Random effect			
Intercept (subjects)	1.020	.242	

**p < .01

comprehension was still significantly related to appreciation, but print exposure was not able to explain any additional variance. The third model, in which we added condition as independent variable, shows that the significant association between perceived comprehension and appreciation is still there and that condition does not explain any additional variance.

Role of print exposure

The mixed models show that print exposure did not have an effect on increases in appreciation within the model. However, to compare the two groups of readers—high print exposure (experience) readers versus low print exposure (inexperienced) readers—we conducted a median split on our sample and investigated the interaction between rereading and reading experience. The results show that depth of appreciation scores were very small for both groups of readers, and both conditions (Figure 3). Both groups of readers even have a negative depth of appreciation score for the manipulated story. The sensitivity index for the experienced readers (–.07) and of the inexperienced readers (.04) indicate that both groups were hardly sensitive to our manipulation. The difference between the two groups of readers is again negligible (.07).

Since we originally used a different measure of reader expertise than Dixon et al. (1993), we wanted to check whether the ART scores and the self-reported reading frequency measure capture similar reader characteristics (i.e., their experience as a reader) by running a correlation analysis. The question we asked participants to measure reading frequency was "how many hours per week do you read fiction for your own pleasure?" The participants could answer on a scale of 1 (or fewer hours) to 10 (or more hours). The results revealed a significant but moderately positive correlation between print exposure and reading frequency, $r = .26$, $p = .01$, indicating that the two measures are related but not very closely. Therefore, we decided to run the series of mixed model analyses again, this time including the independent variable of reading frequency instead of print exposure in the second model. The results can be found in Table 12.
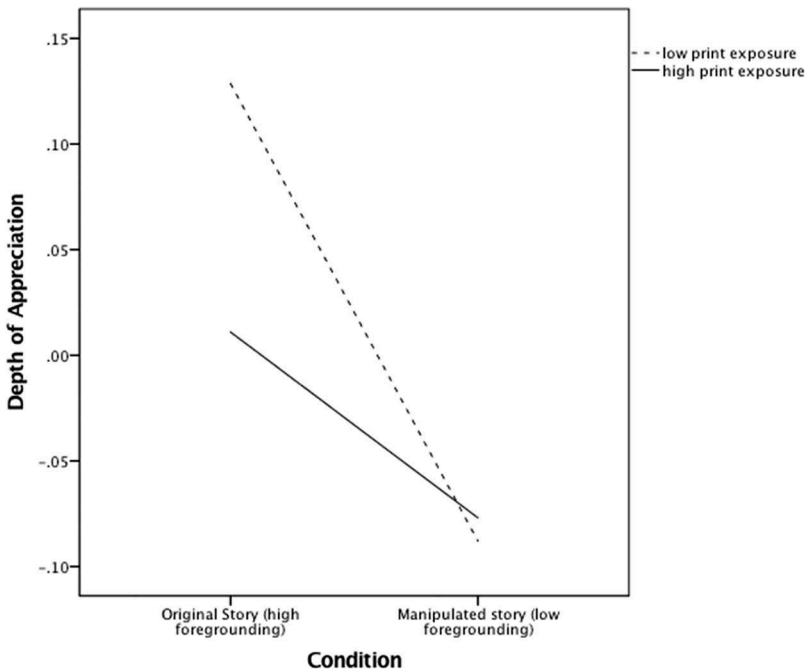

Figure 3. Depth of appreciation by low print exposure and high print exposure readers with and without foregrounding.

Table 12. Fixed Effect Estimates and Variance-Covariance Estimates for Repeated Measures Mixed Models with Reading Frequency Experiment 3.

	Estimator	Standard Error	t(df)
Model 1			
Fixed effect			
Comprehension	.374**	.105	3.57 (94)
Random effect			
Intercept (subjects)	1.026	.236	
Model 2			
Fixed effects			
Comprehension	.422**	.118	3.57 (85)
Reading frequency	−.087	.067	−1.30 (56)
Random effect			
Intercept (subjects)	.963	.225	
Model 3			
Fixed effects			
Comprehension	.413**	.119	3.47 (83)
Reading frequency	−.093	.067	−1.37 (55)
Condition	.229	.315	.73 (56)
Random effect			
Intercept (subjects)	.957	.226	

$**p < .01$

We conducted the first mixed model with appreciation as a time-varying dependent variable and perceived comprehension as a time-varying covariate. We found a small significant positive relation between perceived comprehension and appreciation: As perceived comprehension increases, so does appreciation ($R^2 = .08$). When adding reading frequency as a covariate to the second model, the significant relation with perceived comprehension was still there, but reading frequency did not explain any additional variance. In the third model we added condition as independent variable, and again the significant effect of perceived comprehension was still in place, whereas condition did not
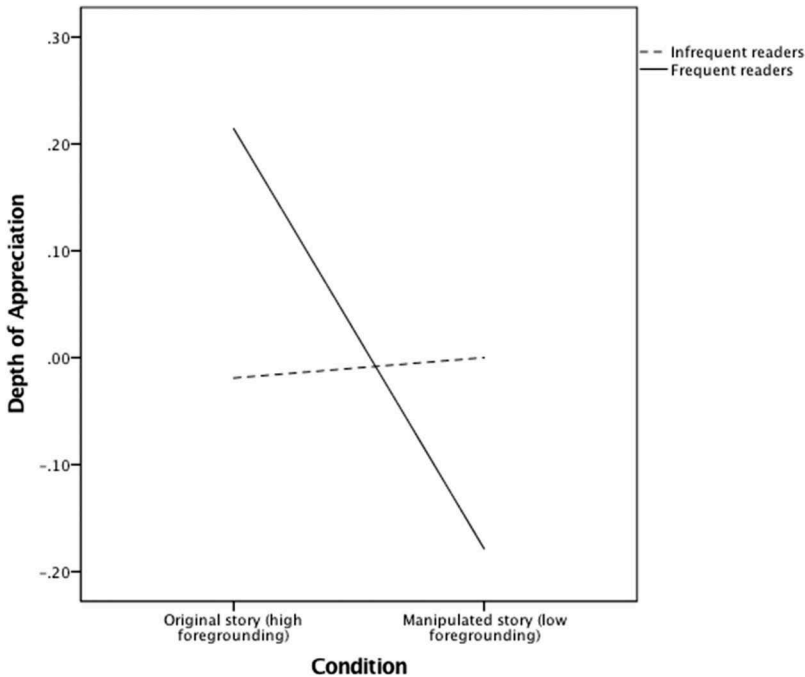

Figure 4. Depth of appreciation by infrequent and frequent readers with and without foregrounding.

explain additional variance. Interestingly enough, when we ran the median split analyses, the results of which can be found in Figure 4, we were able to replicate the pattern of results from Dixon et al.'s original study. Even though reading frequency and print exposure were highly correlated, these analyses reveal that they do seem to measure different phenomena. Nevertheless, the differences again were very small: The sensitivity index for frequent readers was .03 and for the infrequent readers .02. Overall, reading frequency as well as print exposure did not contribute significantly to the linear mixed model in the third study.

Discussion

The results of Experiment 3 show that possible increases in appreciation from first to second reading are not related to levels of literariness in the text but rather to increases in perceived comprehension between first and second reading. Print exposure and reading frequency do not seem to affect increases in appreciation, and thus we were unable to detect a role of the reader, at least, not to the extent we had expected considering the results of Dixon et al. (1993).

General discussion

Previous research shows that rereading literary texts increases appreciation of those texts and that certain literary text features are likely to cause this effect in interaction with an audience of frequent readers. In the three experiments presented here we were mostly unable to replicate these findings. The increases in appreciation we did register did not seem to have been caused by literary style aspects: Changes in responses to our measures could not be linked to predetermined levels of foregrounding of the narratives. Instead, we found that increases in appreciation were related to increases in perceived comprehension. This finding is consistent across texts, manipulations, and samples. It seems, therefore, that appreciation increases over time and that it is grounded in emerging comprehension. Previous research shows that this is due to literariness; the present studies suggest that the emergent effects are not much different from what we might expect to see in response to other nonliterary or less literary texts. Before we elaborate on this proposition, we consider a few alternative suggestions in an attempt to understand why our present results do not seem to concur with those of the other studies we discussed in the introduction.

Foregrounding and emergent literary effect

Why did our participants not respond the way we predicted based on previous research? There are a number of reasons we can think of. Obviously, it is possible to read a literary text and not have a literary reading experience. For example, it may be that the stylistic deviations do affect readers but that they are not aware of it or that our measures were unable to detect them. Moreover, it cannot be ruled out that the theories about the effects of literary devices are invalid. On the other hand, we did see some hints in the data that it may be participants' attitude toward literature that defines the outcome of the treatments we submit them to. For example, in the second study, in which we collected responses of supposedly dedicated readers, (i.e., book club members), we seemed to have hit on an emergent literary effect (other than increased perceived comprehension) that accounted for some additional variance in appreciation between readings in the case of the high foregrounded text. To us, it seems likely that the attitude toward literature of this particular sample affects readers' perception of the type of reward that rereading brings them.

Another alternative explanation pertains to the particular manipulation we applied. We assume that various *forms* of foregrounding may yield distinct *kinds of experienced rewards*. We adjusted levels of foregrounding (Experiments 2 and 3) through systematic rewriting based on the Miall and Kuiken (1994) method of analysis. The advantages, as we saw them, were that foregrounding was changed throughout the text, that the procedure was replicable for other texts, and quantifiable, thus enabling us to test the extent of the manipulation statistically. However, from the perspective of readers, it may be that other forms of

foregrounding are more impactful. It may very well be that specific forms of foregrounding lead to distinct forms of responses and that these are not to be found throughout the text but much more local, for instance in a particular metaphor (cf. Bálint et al., 2016; Koopman, 2016). Future research might investigate hypotheses that are more grounded in such experiences rather than in theories. It may be that these will lead to more narrowly targeted text manipulations than applied in the present studies.

Increased comprehension as literary emergent effect

As to the main finding of the present studies, we propose that the nature of the emerging effects is an emerging comprehension that co-occurs with or causes increases in appreciation. In the Introduction we suggested a few ways in which rereading could contribute to appreciation (i.e., reallocation of cognitive resources, increased processing fluency) processes that are closely related to an increase in comprehension. Research in education and developmental psychology has shown that increased comprehension through rereading can contribute to appreciation in different ways. Levy, Nicholls, and Kohen (1993) found that good and poor readers in grades 3, 4, and 5 gained comprehension through rereading, not because they allocated their cognitive resources to other parts of the texts and filled in gaps in the text from prior memory but by simply becoming more efficient in word recognition and other comprehension processes (p. 321). Rereading, in other words, contributes not only to better comprehension but also to readers improving their reading skills. And this, we assume, in turn can lead to an increase in appreciation. Of course, we cannot infer from our present results that an increase in reader skill occurs through rereading and that this could be related to increases in appreciation. Nevertheless, we believe this is a viable avenue for further research.

Limitations

While discussing the alternative interpretations of our findings, we already stipulated a number of questions left unanswered by the present work. First, our results might hint at the possibility of a causal relationship: A better understanding upon rereading allows readers to appreciate the story more. However, we are unable to make such causal claims based on the results of this study due to the nature of our comprehension measure. We only know that rereading increased readers' self-reported perceived or meta-comprehension, but we do not know whether they actually comprehended the text better upon rereading. Future studies could use implicit measures such as eye movements and reading time as a more objective measure of increased comprehension. As Raney and Rayner (1995, pp. 167–168) suggested, "rereading time might represent a natural measure of comprehension (parts of the text which are better understood should show the largest decrease in reading time)."

As to the methodology used, we have a few reservations. First, the method of text manipulation has its own limitations, often altering not just the target features of a text but also additional features such as text length or text meaning. Second, we propose that the rereading paradigm as an empirical method in itself needs to be investigated. Natural reading is probably different from reading in an experimental setting, and the same goes for *rereading*. As we pointed out in the Introduction, rereading is a phenomenon that is ecologically valid and therefore worth empirical investigation. However, the way in which empirical studies have investigated rereading so far, with help of the rereading paradigm, might differ from naturally occurring rereading in important ways. Dixon et al. proposed that "good literature is processed over an extended period of time" (1993, p. 14). In future studies the extended period might be included in the research design, investigating rereading effects after longer periods of time using, for instance, experience sampling methods (Conner, Tennen, Fleeson, & Barrett, 2009) and other qualitative methods. Finally, it may be argued that in the present studies we are assessing a *retesting* effect rather than a *rereading* effect. Participants knew what kind of

questions to expect after the second reading, which might have influenced how they responded the second time. Of course, with our first study we found that there was no difference between immediate rereading—in which the risk of a retest effect increases—and rereading after a week. Still, even after a week there is some level of validity threat remaining when it comes to retest effects. Studying the topic of rereading entails using a within-subjects repeated measures design, which might affect how participants read and evaluate the second time. One simple option for a randomized between-subject design that might solve these problems is to have one group read a text once and another twice. The advantage of the rereading paradigm is, however, that it allowed us to investigate the development of the responses and the role that changes in comprehension have in this.

Outlook

Comprehension seems an important condition for emergent literary effects. Further research might look into the extent to which this is the case. In addition, other factors might be considered as well, such as *perceived* foregrounding. It may be that the phenomenon of "comprehension as emergent effect" after rereading is common in responses to various genres. However, we propose that it typically leads to increases in appreciation in responses to literary texts. Rereading a text on mathematics, philosophy, or law, for example, may increase understanding but would lead to different outcomes characteristic of those genres. It is important *what* readers understand better. In case of a text on mathematics, it is some form of calculation that readers see better or a notion in a philosophical text that they comprehend better. In a story, it is a different reward that is awaiting them. Future studies might try to sort out in more detail what these rewards are, so as to understand why readers return to certain stories and reread them again and again.

Acknowledgements

Funding

ORCID

Moniek M. Kuijpers ⓘ http://orcid.org/0000-0002-3676-5879

References

Alter, A. L. (2013). The benefits of cognitive disfluency. *Current Directions in Psychological Science*, *22*, 437–442. doi:10.1177/0963721413498894

Atkins, J. (1934). *Literary criticism in antiquity*. Cambridge: Cambridge University Press.

Bálint, K., Hakemulder, F., Kuijpers, M. M., Tan, E. S., & Doicaru, M. M. (2016). Reconceptualizing foregrounding: Identifying response strategies to deviation in absorbing narratives. *Scientific Study of Literature*, *6*, 176–207. doi:10.1075/ssol.6.2.02bal

Berlyne, D. E. (1970). Novelty, complexity, and hedonic liking. *Attention, Perception and Psychophysics*, *8*, 279–286. doi:10.3758/BF03212593

Bierce, A. (1975). *Bij de brug over de owl creek [An occurrence at owl creek bridge]* (B. Den Uyl, Trans.). Amsterdam: Meulenhoff.

Borges, J. L. (2003). Emma Zunz. In *De Aleph en andere verhalen* (B. van de Pol, Trans.). Amsterdam: De Bezige Bij.

Brecht, B. (1976). Kleines organon fur das theater. In *Gesammelte werke* (pp. 659–708). Frankfurt am Main: Suhrkamp.

Conner, T. S., Tennen, H., Fleeson, W., & Barrett, L. F. (2009). Experience sampling methods: A modern idiographic approach to personality research. *Social and Personality Psychology Compass*, 3, 292–313. doi:10.1111/j.1751-9004.2009.00170.x

Cutting, J. (2006). The mere exposure effect and aesthetic preference. In P. Locher, C. Martindale, & L. Dorfman (Eds.), *New directions in aesthetics, creativity and the arts* (pp. 33–46). Amityville, NY: Baywood.

Dixon, P., Bortolussi, M., Twilley, L., & Leung, A. (1993). Literary processing and interpretation: Toward empirical foundations. *Poetics*, 22, 5–33. doi:10.1016/0304-422X(93)90018

Farkas, A. (2002). Prototypicality-effect in surrealist paintings. *Empirical Studies of the Arts*, 20, 127–136. doi:10.2190/UD7Y-GN8P-Q0EV-Q13J

Green, M., Kass, S., Carrey, J., Herzig, B., Feeney, R., & Sabini, J. (2008). Transportation across media: Repeated exposure to print and film. *Media Psychology*, 11, 512–539. doi:10.1080/15213260802492000

Hakemulder, F., & Van Peer, W. (2015). Empirical stylistics. In V. Sotirova (Ed.), *Bloomsbury companion to stylistics* (pp. 189–207). Oxford: Bloomsbury.

Hakemulder, J. (2008). The more you see, the more you get: How spectators' use their limited capacity for attention in responses to formal aspects of film. In J. Auracher & W. Van Peer (Eds.), *New beginnings in literary studies* (pp. 332–352). Newcastle: Cambridge Scholars Publishing.

Hakemulder, J. F. (2004). Foregrounding and its effect on readers. *Discourse Processes*, 38, 193–218. doi:10.1207/s15326950dp3802_3

Hunt, R., & Vipond, D. (1985). Crash-testing a transactional model of literary reading. *Reader: Essays in Reader Oriented Theory, Criticism, and Pedagogy*, 14, 23–39.

Koopman, E. M. (2015). Empathic reactions after reading: The role of genre, personal factors and affective responses. *Poetics*, 50, 62–79. doi:10.1016/j.poetic.2015.02.008

Koopman, E. (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

Kuchinke, L., Trapp, S., Jacobs, A. M., & Leder, H. (2009). Pupillary responses in art appreciation: Effects of aesthetic emotions. *Psychology of Aesthetics, Creativity, and the Arts*, 3, 156–163. doi:10.1037/a0014464

Kuijpers, M. M. (2014). *Absorbing stories: The effects of textual devices on absorption and evaluative responses* (Unpublished PhD thesis). Utrecht University.

Leavitt, J. D., & Christenfeld, N. J. S. (2011). Story spoilers don't spoil stories. *Psychological Science*, 22, 1152–1154. doi:10.1177/0956797611417007

Leder, H., & Nadal, M. (2014). 10 years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode - Developments and challenges in empirical aesthetics. *British Journal of Psychology*, 105, 443–464. doi:10.1111/bjop.12084

Levy, B. A., Nicholls, A., & Kohen, D. (1993). Repeated readings: Process benefits for good and poor readers. *Journal of Experimental Child Psychology*, 56, 303–327. doi:10.1006/jecp.1993.1037

Martindale, C., & Moore, K. (1988). Priming, prototypicality, and preference. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 661–670. doi:10.1037/0096-1523.14.4.661

Menninghaus, W., Bohrn, I. C., Knoop, C. A., Kotz, S. A., Schlotz, W., & Jacobs, A. M. (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*, 143, 48–60. doi:10.1016/j.cognition.2015.05.026

Miall, D. S., & Kuiken, D. (1994). Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics*, 22, 389–407. doi:10.1016/0304-422X(94)00011-5

Millis, K. K., Simon, S., & Tenbroek, N. S. (1998). Resource allocation during the rereading of scientific texts. *Memory & Cognition*, 26, 232–246. doi:10.3758/BF03201136

Mol, S., & Bus, A. (2011). To read or not to read: A meta-analysis of print exposure from infancy to adulthood. *Psychological Bulletin*, 137, 267–296. doi:10.1037/a0021890

Mukarovsky, J. (1964). Standard language and poetic language. In P. L. Garvin (Ed.), *A Prague school reader on esthetics, literary structure, and style* (pp. 17–30). Washington, DC: Georgetown University Press.

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology*, 49, 151–172. doi:10.1037/1196-1961.49.2.151

Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition*, 28, 1004–1010. doi:10.3758/BF03209348

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8, 364–382. doi:10.1207/s15327957pspr0804_3

Rushdie, S. (2003). *Middernachtskinderen [Midnight's children]* (M. Schuchart, Trans.) (11th ed.). Amsterdam: Pandora, Contact BV.

Sanford, A., Sanford, A., Molle, J., & Emmott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes*, 24, 109–130. doi:10.1207/s15326950dp4202_2

Sanford, A. J., & Emmott, C. (2012). *Mind, brain and narrative*. Cambridge: Cambridge University Press.

Schmidt, S. J. (1982). *Foundations for the empirical study of literature*. Hamburg: Buske.

Shklovsky, V. (1965). Art as technique. In L. T. Lemon & M. J. Reis (Eds. & Trans.), *Russian formalist criticism: Four essays* (pp. 3–24). Lincoln, NE: University of Nebraska Press.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 24, 402–433. doi:10.2307/747605

Stempel, W.-D. (Ed) (1972). *Texte der russischen Formalisten II*. Munich: Fink.

Van Peer, W. (1986). Pulp and purpose: Stylistic analysis as an aid to a theory of texts. In T. D'Haen (Ed.), *Linguistics and the study of literature* (pp. 268–286). Amsterdam: Rodopi.

Wassiliwizky, E., Koelsch, S., Wagner, V., Jacobsen, T., & Menninghaus, W. (2017). The emotional power of poetry: Neural circuitry, psychophysiology and compositional principles. *Social Cognitive and Affective Neuroscience*, 12, 1229–1240. doi:10.1093/scan/nsx069

West, B. T., Welch, K. B., & Galecki, A. T. (2015). *Linear mixed models: A practical guide using statistical software*. Boca Raton, FL: CRC Press.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, 81, 989–1000. doi:10.1037/0022-3514.81.6.989

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9, 1–27. doi:10.1037/h0025848

Zyngier, S., van Peer, W., & Hakemulder, F. (2007). Complexity and foregrounding. In the eye of the beholder? *Poetics Today*, 28, 653–682. doi:10.1215/03335372-2007-011