



Literature and readers' empathy: A qualitative text manipulation study

Language and Literature
2017, Vol. 26(2) 137–152

© The Author(s) 2017

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0963947017704729

journals.sagepub.com/home/lal



Anežka Kuzmičová

Stockholm University, Sweden

Anne Mangen

University of Stavanger, Norway

Hildegunn Støle

University of Stavanger, Norway

Anne Charlotte Begnum

University of Stavanger, Norway

Abstract

The alleged crisis of the humanities is currently fueling renewed interest in the affective benefits of literary reading. Several quantitative studies have shown a positive correlation between literary reading and empathy. However, the literary nature of the stimuli used in these studies has not been defined at a more detailed, stylistic level. In order to explore the stylistic underpinnings of the hypothesized link between literariness and empathy, we conducted a qualitative experiment in which the degree of stylistic foregrounding was manipulated. Subjects ($N = 37$) read versions of Katherine Mansfield's "The Fly," a short story rich in foregrounding, while marking striking and evocative passages of their choosing. Afterwards, they were asked to select three markings and elaborate on their experiences in writing. One group read the original story, while the other read a 'non-literary' version, produced by an established author of suspense fiction for young adults, where stylistic foregrounding was reduced. We found that the non-literary version elicited significantly more ($p < 0.01$) explicitly empathic responses than the original story. This finding stands in contradiction to widely accepted assumptions in recent research, but can be assimilated

Corresponding author:

Anežka Kuzmičová, Department of Culture and Aesthetics, Stockholm University, SE-10691, Stockholm, Sweden.

Email: anezka.kuzmicova@littvet.su.se

in alternative models of literariness and affect in literary reading. We present an analysis of the data with a view to offering more than one interpretation of the observed effects of stylistic foregrounding.

Keywords

Reader response, empathy, literary fiction, foregrounding, qualitative methods

I. Literature and empathy

Experimental research (Kidd and Castano, 2013a) has recently indicated that literary reading may be positively correlated with increased empathy and/or affective theory of mind. Based on these findings, it is more and more frequently suggested that literary fiction fosters interpersonal skills and pro-social behavior, and that it does so to a greater extent than both non-fiction and so-called popular fiction. The long-term as well as short-term effects of literary reading on empathy are currently being investigated and we review these findings below.

1.1 Long-term effects

In reader response experiments, readers' long-term exposure to literature is often measured using the Author Recognition Test (ART) (Mar et al., 2006; Stanovich and West, 1989). The ART is a checklist of names, some of which belong to well-known writers of either literary or popular fiction. Respondents are asked to select items that they recognize as writers' names, scoring points for all correct answers while points are deducted for incorrect selections. Relying on the ART, Kidd and Castano (2013a) observed a positive correlation between literary reading and theory of mind (ToM), i.e. the ability to accurately identify the mental states of other people. In their series of experiments, which received considerable media coverage (e.g. Belluck, 2013), theory of mind was primarily measured with the Reading the Mind in the Eyes Test (RMET) (Baron-Cohen et al., 2001). The RMET consists of photographs portraying human eyes. Each photograph is meant to express a discrete emotion and respondents are asked to select the correct emotion in a multiple-choice setup. While the concept of theory of mind is usually understood to comprise the assessment of both affective (emotions) and cognitive (intentions, beliefs) mental states, additional ToM tests (Shamay-Tsoory and Aharon-Peretz, 2007) administered by Kidd and Castano suggested that the observed effects were limited to affective theory of mind exclusively.

Varieties of the ART have also been used in order to study the long-term effects of distinct genres on readers' affective ToM. For instance, Mar et al. (2006, 2009) and Djikic et al. (2013) found that an overall preference for fiction over expository non-fiction was also associated with higher RMET scores. Although a distinction between fiction and expository non-fiction is not the same as a distinction between literary and non-literary writing, it is possible that there are overlaps between these categories. The findings can thus be considered indirectly relevant to the question of empathy as a long-term effect of literary reading. In a study by Koopman (2015), long-term exposure to

literature was found to correlate with higher scores on a self-report measure of empathic understanding toward individuals in distress, administered as part of a reader response experiment. The self-report measure consisted of a series of first-person statements (e.g. “I feel understanding for people who are depressed;” “I can imagine it must be horrible to be depressed”) that were rated by the subjects on a seven-point Likert scale.

1.2 Short-term effects

However, the main purpose of Koopman’s (2015) experiment was to examine the potential of literary and non-literary texts for inducing immediate empathic feelings and pro-social conduct. Each participant read two texts belonging to one of the following genres: expository (non-narrative) text; life (non-fiction) narrative; or literary narrative. The common topic of the texts, which were read a week apart, was either depression or grief. In addition to administering the self-report measure of empathic understanding, Koopman also investigated whether the empathy hypothetically induced by textual stimuli had real-life consequences in terms of donating behavior. She observed that in the week following the first session, subjects who had read about depression in the life narrative condition, specifically, donated more to a related charity than any of the other groups. There was no evidence of literature’s superiority over other genres in inducing empathy or pro-social conduct in the short term.

Kidd and Castano (2013a) also investigated whether the positive effects of literature on empathy might be observed in the short term. Therefore, their series included experimental designs wherein each subject read a short story considered by the experimenters to be either literary or popular. After reading the story, the participants were tested for theory of mind. Increased RMET scores were indeed observed in the literary condition. By contrast, Djikic et al. (2013) found no effects of genre on either the RMET or self-report empathy measures in a similar design comparing critically acclaimed literary fiction to non-fiction. In another study, Bal and Veltkamp (2013) found that fiction increased subjects’ general capacity for empathy as measured by a self-report scale, but only under the condition of high emotional transportation¹ into the story. Lastly, Johnson (2013) found that, via transportation and empathy, a literary narrative was capable of reducing out-group prejudice, but no genre control was used in this design.

Overall, the evidence that literary fiction elicits more empathy in readers than other genres is inconclusive. Moreover, the studies referred to above refrain from describing the experimental stimuli and observed between-genre differences in terms of a key variable, i.e. style. For instance, Koopman (2015) was constrained by the nature of her experiment to select varied texts about grief and depression rather than being able to control closely for stylistic nuance between the different genre conditions. Kidd and Castano’s literary and popular stories were sampled without consideration of their stylistic properties, resulting in a diverse mix of critically acclaimed fiction (Kidd and Castano, 2013b) and stories from an anthology labeled and marketed as ‘popular’. Across all the above-mentioned studies, it is impossible to determine specifically which of the many stylistic features characteristic of literary fiction (e.g. Miall, 2006) ought to be hypothesized as more likely than others to elicit empathy. More importantly, it is impossible to determine whether the observed effects were really due to the distinctiveness of a given

genre (literature, life narrative) rather than being the result of incidental differences between the stimuli in plot structure, number of characters, narrative perspective, and so forth.

2. The qualitative text manipulation study

2.1 Methodological and theoretical background

To attempt a more nuanced, stylistically informed account of the hypothetical nexus of literature and empathy, we used an experimental design in which we manipulated a literary text to construct an alternative non-literary stimulus, instead of sampling two different unaltered texts. Our literary stimulus was selected on the basis of the presence of stylistic *foregrounding*, i.e. its potential for defamiliarization through the use of deviant linguistic devices.

A notion originating in the early 20th century theoretical traditions of formalism and structuralism, the systematic use of foregrounding has been repeatedly proposed as one of the hallmarks of literary texts (for a review, see e.g. Gavins, 2014). While the initial explanations of the defamiliarizing effect of foregrounding remained somewhat unarticulated in the suggestion that foregrounding “(removes) the automatism of perception” (Shklovsky, 1988 [1925]: 27), later work in empirical stylistics and literary studies (van Peer, 1986) advanced the foregrounding framework by making it operational in experimental setups. Perhaps most notably, reader response experiments (Fialho, 2007; Miall and Kuiken, 1994) have yielded evidence of foregrounding prompting the so-called defamiliarization–feeling–refamiliarization cycle, wherein readers come to ponder a foregrounded expression, experience novel feelings and worldviews in response to it, and move on to align these with their previous cognitive-affective grasp of the text as well as the world beyond the text. Literary writing is thus understood to elicit more emotions than non-literary writing (see also Miall, 2011).

In our experiment, an alternative non-literary version of the original stimulus text was created, in which foregrounding was reduced. Other features such as plot structure, number of characters, and narrative perspective were preserved. Text manipulations of this kind are an established method in reader response experiments (Bortolussi and Dixon, 2003) as they enable measuring the effects of relatively circumscribed textual features, e.g. single word units. One of the pitfalls of the method, however, is that subjects in the manipulated condition are exposed to an experimenter-created artificial text that has never been part of the world outside the lab. In order to preserve the narrative quality of the manipulated text, thus ensuring (as far as it is possible to do so) that the two texts were recognized as professional prose, an established writer of popular fiction was commissioned to do the text manipulation in our study.

Manipulation studies of foregrounding have previously been carried out by Hakemulder (2000, 2004, 2008). In one of these studies (Hakemulder, 2004), it was found that a literary text relating to the topic of immigration induced more positive personal attitudes towards immigrants as compared to a manipulated control stimulus where foregrounding had been reduced. The manipulation involved a shortening and simplification of sentence structure, replacement of stylistic figures such as metaphor or irony

with more literal expressions, and a shift from a “baroque” to a “sober” (Hakemulder, 2004: 200–201) style more generally. For example, the expression “[wives] had been grilled by reasonable, doing-their-job officials about the length of and distinguishing moles upon their husbands’ genitalia” was replaced with “[they] had been questioned about intimate details about their husbands.” One could speculate that the observed attitudinal effect, measured through a series of statements concerning immigrants’ life conditions rated on a seven-point Likert scale, may have been mediated by empathy or empathy-related responses to the text. As Hakemulder refrains from investigating the subjects’ first-person experiences in further detail, however, it is impossible to determine how the subjects really felt about the immigrant protagonists while they were reading. The same concern applies to all the quantitative studies reviewed in section 1.

Our study attempts to redress this imbalance between qualitative and quantitative methods by considering participants’ subjective empathic responses to texts in a qualitative text manipulation experiment. Following a paradigm introduced by Sikora et al. (2011), we asked our subjects to mark any text passages that they found particularly striking or evocative in the course of reading. As a next step, the subjects were asked to select three of their markings and elaborate freely in writing on *how* the passages were striking or evocative to them. In light of Hakemulder’s (2004), Kidd and Castano’s (2013a), and Djikic et al.’s (2013) findings, we hypothesized that the original literary text would yield more spontaneously empathic elaborations than the manipulated non-literary version. We also expected to find a quantitative and qualitative difference in passage markings, with a higher number of markings in the literary stimulus and a different selection of passages between the two conditions.

2.2 Stimulus and manipulation

The original literary text chosen for our study was “The Fly,” a short story by the modernist author Katherine Mansfield (1923). As the study was carried out in Norway, a grammatically updated 1950s translation into Norwegian (Bokmål Standard; Mansfield, 1950) was used. Mansfield’s prose can generally be characterized as highly emotion-laden (Kuivalainen, 2009). Due to their everyday themes and moderate foregrounding compared to other Anglophone Modernist writers (e.g. Woolf, Joyce, Pound, or Eliot), Mansfield’s texts have been used extensively and productively in reader response experiments (Fialho, 2012; Hakemulder et al., 2016; Miall, 2006).

“The Fly” describes a brief and seemingly eventless meeting between two old acquaintances: a factory director, referred to solely as “the boss,” and a retired businessman named Woodifield. The narrative point of view, marked by free indirect discourse, largely gravitates toward that of the director. Inadvertently reminded by Woodifield about the death of his only son in WWI, the director releases his visitor in order to briefly contemplate his suppressed grief, before seeking relief in tormenting a fly to death. The narrative style of “The Fly” was relatively novel at the time of publication, in that key emotions are indirectly mediated through small talk and framed by laconically mundane actions such as the unlocking of a cupboard. The latter, in combination with anaphora and other deictic markers (e.g. spatial adverbs), hypothetically facilitate the reader’s sense of “being there” (Kuzmičová, 2012) and make the story relatively accessible.

The manipulated, non-literary version of “The Fly” was prepared in collaboration with Terje Torkildsen (2014), an award-winning Norwegian author of suspense fiction for young adult non-readers. Throughout the story, four main types of manipulation were performed in order to match the popular style with which Torkildsen’s readers are familiar. Firstly, a number of figurative expressions were removed or replaced by more literal expressions (backgrounding). Secondly, a number of indeterminate descriptive expressions were replaced by expressions at or closer to a basic level (see Rosch, 1978) of determinacy (specification). Thirdly, archaic and/or formal grammatical and semantic features were replaced by contemporary and/or more informal equivalents (leveling). Fourthly, a number of complex paratactic structures were broken down into simpler structures (parceling). The manipulations were uniformly distributed throughout the text. As a consequence of the manipulations, the word count of the non-literary version was 9% lower than that of its unaltered counterpart.

Examples of the manipulation procedure are shown below. However, due to differences in innate variability between Norwegian and English, our translations of leveling (especially modernization) are only rough approximations of the instances of leveling used in the Norwegian stimuli.

Manipulation examples:

B = backgrounding; L = leveling; P = parceling; S = specification

[1]

Literary condition

As a matter of fact he was proud of his room; he liked to have it admired, especially by old Woodifield. It gave him a feeling of deep, solid satisfaction to be planted there in the midst of it [...]

Non-literary condition

He was proud of his office [L]; he liked when people admired it [L], especially old Woodifield. It gave him a feeling of power [S] when he sat there [B] in his chair [S] [...]

[2]

Literary condition

But he did not draw old Woodifield’s attention to the photograph over the table of a grave-looking boy in uniform standing in one of those spectral photographers’ parks with photographers’ storm-clouds behind him. It was not new. It had been there for over six years.

Non-literary condition

But he did not point at [L, S] the photo [L] over the table. The one [P] that showed a grave-looking boy in uniform. Its background clearly revealed that it had been taken at a photographer’s [B]. The photo was not new. It had been hanging there [S] for over six years.

2.3 Full experiment

Thirty-seven subjects (31 females) were recruited from among a cohort of Norwegian language and literature teacher-training undergraduates at a Norwegian university. Two additional subjects were excluded from the sample because their qualitative data sets were incomplete. In a between-subject design, one group read the slightly grammatically modernized Norwegian translation, by Emil Boyson, of Mansfield's "The Fly" (the literary condition, 17 subjects, 15 females). Another group (20 subjects, 16 females) read the manipulated version created by Torkildsen (the non-literary condition). During reading, they were asked to mark with a pen any passages that they found particularly striking or evocative. Once this task was finished, they were asked to select three of their markings and elaborate in writing on how they had experienced the passages as striking or evocative. The elaborations were written on a computer. In the same session, the subjects rated their immediate reading experiences on a number of variables using a computerized post-process questionnaire that was largely based on extant measures of transportation (Kuijpers et al., 2014) and narrative engagement (Busselle and Bilandzic, 2009). After that, they also completed the RMET in a clinically piloted Norwegian translation (Sommerfeldt and Skårderud, 2008).

The qualitative study was part of a larger experimental design² and took place during the latter of two experimental sessions. During Session 1, preceding the qualitative study by three weeks, the subjects' baseline theory of mind score had been collected in a first trial of the RMET. The subjects had also completed a questionnaire targeting their general attitudes to literature, largely adapted from Miall and Kuiken (1995), and a demographics and reading habits questionnaire partly adapted from Acheson et al. (2008).³ In addition, they had completed a reading comprehension task adapted from PISA 2000 (OECD, 2002). The attitudes questionnaire administered in Session 1 and the post-process questionnaire administered in Session 2 both included items relevant to empathy during reading.

2.4 Elaborations and coding

With 37 subjects, each producing three elaborations, in the qualitative design, we collected a total of 111 elaborations: 51 in the literary condition and 60 in the non-literary condition. In view of the hypothesized association between literariness and affective theory of mind, the elaborations were categorically coded for explicit markers of empathic response. A number of further, partially nested categories emerged from the content of the elaborations as shown in Table 1 and Coding examples 3–6 below.

Coding examples:

[3]

Item 35, subject 12, non-literary condition; ToM

This is a clear illustration of how deeply affected one can be [*Gen*] by losing a loved one. As reader I get to feel [*Emp*] the father's grief [*Spec*].

[4]

Item 2, subject 1, non-literary condition; ToM

This is quite evocative, he is probably [*Mod*] sad [*Spec*] after the visit and after being reminded of his sorrow [*Spec*] for his deceased son.

[5]

Item 106, subject 36, literary condition; ToM and Non-ToM

I get a sense of what it's like in his office [*Imag*]. These are nice descriptions that make it easy to put myself in [*Emp*] the character's position. Also, the boss is described in a sort of unexpected way [*Styl*] compared to what you might imagine when you think of a "boss."

[6]

Item 77, subject 26, literary condition; Non-ToM

New suspense begins to build up [*Plot*] at this point in the text. It's because this is where the fly is mentioned for the first time. The Fly is also the title of the story, so one gets curious about what the title is meant to convey.

In the literary condition, elaborations tended to cluster around the foregrounded expressions that were revised or removed in the non-literary condition. However, no significant difference in number of spontaneous markings was observed between the two conditions. In most of the elaborations, subjects reported on their evolving reactions in the course of reading, even though the elaborations were written afterwards. Only a few of the elaborations overtly interpreted the marked passages in light of what the subjects had learned from the text as a whole.

Table 1. Coding categories.**Theory of Mind-related qualities**

ToM: item refers to story character's affective and/or cognitive mental state.

Spec: character is attributed a specific emotion, intention, belief, or other mental state.

Mod: epistemic modality is used in attributing a specific mental state to character.

Gen: character's specific mental state is generalized based on subject's real-world knowledge.

Emp: item explicitly refers to subject's first-person experience of empathy with character's mental state (specific or not).

Other (non-ToM) qualities

Non-ToM: item refers to other story qualities unrelated to theory of mind.

Plot: item refers to subject's suspense or surprise in relation to plot.

Imag: item refers to subject's experience of sensory mental imagery.

Styl: stylistic features are explicitly appreciated and/or described.

2.5 Results

Generally, a majority of the elaborations referred to some type of ToM response as the main reason for marking the passage in question. This can be explained with reference to the highly emotional topic of the story (child loss) and its narrative technique, wherein the boss's grief is largely implied rather than explicitly expressed (see also Kuivalainen, 2009). It is important to note that this technique was preserved also in the manipulated stimulus, where added specifications mainly concerned concrete objects and where mental states were made more specific only sporadically and only in relation to other, less central emotions and motives (see the manipulation examples in section 2.2 above). The literary vs. non-literary condition had no effect on elaboration length.

Given the relatively small size of our categorical data set, a standard chi-square test could not be performed to calculate statistically significant associations for most pairs of key variables. Fisher's exact test (two-tailed p -values) was used instead to identify statistically significant associations among the variables. As a main finding, we observed a robust effect of the literariness variable on readers' empathic responses. However, this effect was contrary to what we expected based on the research reviewed above, with the non-literary condition eliciting significantly more explicitly empathic responses than the literary condition (ToM Emp in Table 1). The effect was present when all elaborations were treated indiscriminately ($p = 0.0001$) as well as when individual subjects' elaborations were treated separately in clusters of three ($p = 0.0055$). This means that the observed effect should not be attributed to pre-existing between-subject differences in personal response preferences alone, i.e. to an idiosyncratic tendency in a small subset of our subjects, incidentally assigned to the non-literary condition, for explicitly evaluating fictional stimuli in terms of empathy. Fifty percent of the males (3 subjects) and 35% of the females (11 subjects) in our sample made at least one explicit reference to first-person empathy. Although the size of the sample prevents any generalization in this matter, we can conclude that contrary to common expectations (see e.g. Mar et al., 2009), we did not find our male subjects to empathize any less than our female subjects.

Additional findings concern the proportion of ToM vs. non-ToM responses more generally. The non-literary condition yielded significantly more elaborations concerned with ToM exclusively. It also yielded more ToM responses overall, but this effect was not statistically significant. Elaborations in the literary condition, on the other hand, involved significantly more references to non-ToM story qualities, mentioned either exclusively or in combination with ToM, and significantly more references to style. Style was the most frequent non-ToM quality mentioned in the literary condition, whereas plot was the most frequent non-ToM quality in the non-literary condition. However, none of these additional effects persisted after the corpus of elaborations was broken down according to individual respondents. This means that they were somewhat more likely to be artifacts of pre-existing individual differences rather than emergent effects of the literary vs. non-literary condition.

Relative category frequencies (in %) and effect sizes (two-tailed p -values), as calculated per the aggregate corpus of elaborations (Table 2) and per individual respondents (Table 3), are shown below. Statistically significant effects ($p < 0.05$) are marked with an asterisk.

Table 2. Relative category frequencies and effect sizes: aggregate corpus.

	Literary (%)	Non-literary (%)	Effect size (<i>p</i> -value)
ToM overall	53	68	0.1192
ToM exclusively	24	58	0.0003*
– <i>Spec</i>	19	57	0.0565
– <i>Mod</i>	10	10	1.0000
– <i>Gen</i>	02	10	0.1218
– <i>Emp</i>	04	35	0.0001*
Non-ToM overall	76	42	0.0003*
Non-ToM exclusively	47	32	0.1192
– <i>Plot</i>	31	27	0.6755
– <i>Imag</i>	22	08	0.0599
– <i>Styl</i>	37	18	0.0324*

ToM: theory of mind; *Spec*: specific emotion, intention, belief, or other mental state; *Mod*: epistemic modality; *Gen*: character's specific mental state is generalized based on subject's real-world knowledge; *Emp*: subject's experience of empathy with character's mental state; *Plot*: subject's suspense or surprise in relation to plot; *Imag*: subject's experience of sensory mental imagery; *Styl*: stylistic features are explicitly appreciated and/or described.

Table 3. Relative category frequencies and effect sizes: individual respondents.

	Literary (%)	Non-literary (%)	Effect size (<i>p</i> -value)
ToM overall	65	85	0.2502
ToM exclusively	12	20	0.6655
– <i>Spec</i>	53	85	0.0689
– <i>Mod</i>	29	20	0.7034
– <i>Gen</i>	06	25	0.1886
– <i>Emp</i>	12	60	0.0055*
Non-ToM overall	88	80	0.6655
Non-ToM exclusively	35	15	0.2502
– <i>Plot</i>	47	55	0.7459
– <i>Imag</i>	47	20	0.1575
– <i>Styl</i>	65	45	0.3248

ToM: theory of mind; *Spec*: specific emotion, intention, belief, or other mental state; *Mod*: epistemic modality; *Gen*: character's specific mental state is generalized based on subject's real-world knowledge; *Emp*: subject's experience of empathy with character's mental state; *Plot*: subject's suspense or surprise in relation to plot; *Imag*: subject's experience of sensory mental imagery; *Styl*: stylistic features are explicitly appreciated and/or described.

One possible interpretation of the above results is that the literary condition in fact elicited the same amount of empathy as (or even more than) the non-literary condition, but that this empathy was downplayed in the literary items due to subjects choosing to elaborate on other salient qualities of their experiences instead. This interpretation presupposes a within-item trade-off between empathy on the one hand and reference to

non-ToM qualities on the other. However, we found no significant association in the ToM subset between empathy and exclusive ToM focus ($p = 0.2806$). The interpretation was thus rejected.

The main findings of our qualitative manipulation study seem to contradict recent reports (Kidd and Castano, 2013a) that literary fiction is better suited than other genres for eliciting empathy. The additional quantitative measures administered alongside the qualitative study enabled us to control for potential confounds, e.g. interference with individual subjects' literary reading habits and general attitudes to literature. Such interference would suggest that the subjects in our non-literary condition might have been more apt to report empathy because they simply happened to be more avid readers, and thus supposedly better empathizers in the long term. However, neither the reading habits nor the general attitudes questionnaire scores from Session 1 confirmed this hypothesis. Furthermore, interference with subjects' individual empathy dispositions irrespective of long-term exposure or attitude to literature was ruled out on the basis of the RMET scores (Sessions 1 and 2), which in contrast to previous findings (Kidd and Castano, 2013a) showed no significant association with literariness in either direction. In sum, the outcomes of our quantitative measurements indicate that the qualitative differences observed between the literary and non-literary conditions in Session 2 were likely effects proper of the manipulated text variable of literariness.

2.6 Discussion

Interestingly, there was no association between the qualitative findings and self-reported empathy ratings provided in the post-process questionnaire at the end of Session 2. Nor did we find any association with transportation ratings provided in the same questionnaire. A positive association between self-reported transportation and empathic response would raise the possibility that the stimulus in the literary condition did not afford transported reading experience, thus impeding empathy. Such interpretation of the data would be in line with previous findings concerning transportation as a predictor of empathy-related response to narrative (Bal and Veltkamp, 2013; Johnson, 2013). However, it should be noted that the post-process questionnaire was not successfully validated before or within the experiment as an internally consistent psychometric instrument. Although it was adapted from measures that had been previously validated in other language environments, its present outcomes thus have limited reliability.

Assuming that the observed effects indeed were due to the manipulated text variable of literariness, our findings run counter to common expectations but can be accommodated in another theoretical framework: the framework of aesthetic distance. According to this framework, aesthetically marked stimuli are experienced as if from a greater "distance," i.e. in partial awareness of one's pre-existing concerns *as well as* of the fictional world's artificial nature (Cupchik, 2002; Cupchik et al., 1998). This broader affective background is assumed to be bracketed in the reception of more popular (sometimes tellingly labeled "escapist") cultural artifacts.

The aesthetic distance framework agrees with the foregrounding framework inasmuch as distinctly literary stimuli are proposed to implicate the reader to a greater extent. What Miall and Kuiken call refamiliarization, an act of harvesting "[personal memories,

world knowledge] having similar affective connotations” (Miall and Kuiken, 1994; our parenthesis), is akin to the self-oriented perspective entailed by aesthetic distance. The key difference here is that the framework of aesthetic distance as advanced by Cupchik et al. (1998) does not associate foregrounding or aesthetic distance with greater intensity across the full range of emotions. It merely suggests that foregrounding elicits a different set of emotions, allowing for the possibility that some types of emotions, including distinctly empathic ones, could be afforded more generously in less foregrounded texts.

Our initial hypothesis was based upon some of the more recent research discussed in section 1; it relied on an association between readers’ processing effort (e.g. Kidd and Castano, 2013a) and affective theory of mind. The findings of these studies suggest that a popular stimulus relatively lacking foregrounded features would be expected to elicit a more automatic and emotionally shallow, i.e. less empathic, reading. Meanwhile, the alternative framework of aesthetic distance suggests that affective theory of mind may have been equally activated in both our literary and non-literary conditions, but with different emotional outcomes. In a reader response experiment using sampled stimuli, for instance, Cupchik et al. (1998) had subjects report whether they experienced any emotions throughout narrative passages of varied complexity. The subjects also reported whether their emotions were felt to correspond to those of the story protagonist (“fresh emotions”) or whether they rather reflected the subjects’ own past life episodes as reactivated in memory by the narrative stimulus (“emotional memories” distanced from the specific emotions of the protagonist). Cupchik et al. found that more complex passages elicited significantly more of the aesthetically distanced emotional memories compared to more straightforward, descriptive passages, which predominantly elicited fresh emotions, i.e. empathy for the protagonist.⁴

Cupchik et al.’s (1998) findings are consistent with the findings of our qualitative study and suggest that the relative lack of explicitly empathic responses in the literary condition may have been due to aesthetic distance. Applying Cupchik et al.’s account of the different types of emotional response to our study, it could be suggested that both our experimental conditions (literary and non-literary) tapped into affective theory of mind as measurable by the empathy-related items in our post-process questionnaire. Indeed, an item reading “At important moments in the narrative, I could feel the emotions the characters felt,” for instance, received the same average rating (3.8 out of 6 Likert points) across conditions. Moreover, both conditions prompted some transportation-related feelings (e.g. “While reading I was completely immersed in the story;” rated 3.8 in the literary condition, 4.1 in the non-literary condition), which need not always stand in contradiction to aesthetic distance (Cupchik, 2002: 156). Yet only the literary condition elicited a self-oriented affective response supervening empathy with the story protagonist. This latter type of response, i.e. subjects briefly recalling emotive autobiographic memories, falls outside traditional norms of literary analysis in the academia because it fails to advance shareable interpretations of a text as a whole (see also Fialho et al., 2011), hence its lack in subjects’ elaborations.

Alternatively, another distance-based explanation avails, relying even more heavily on the salience of the norms of literary study. According to this explanation, subjects in the literary condition intuitively recognized the highly foregrounded style of the original stimulus as that typical of their academic literary assignments. The wording of their

elaborations was then more or less automatically adjusted to classroom discourse, wherein empathy and other first-person affects tend to be downplayed to the benefit of a more distanced, impersonal analysis. The subjects in the non-literary condition, on the other hand, may have felt more freedom to express empathy due to the stimulus' closer alignment to popular fiction, and thus perhaps treated it as a more leisurely read. In agreement with this interpretation, Cupchik et al.'s (1998) study showed that aesthetic distance is easily manipulated through introducing a particular notion of how, and in what frame of mind, a given story should be read. In addition, the data from the attitudes questionnaire administered in Session 1 reveal some tendency toward socially desirable outcome: e.g. a vast majority of our subjects reported that they valued literature highly, while a full 30% of them disagreed to enjoying reading in their spare time. Thus, it is not unlikely that our qualitative data partly reflects differences in stylistically triggered social norms (see e.g. Allington, 2011) rather than portraying the subjects' entirely private aesthetic reactions.

3. Conclusion

Through either interpretation, our qualitative study fails to confirm the widespread hypothesis that a literary style elicits more empathy than a more popular one, suggesting instead that it elicits a more aesthetically distanced reading. It is important to note, however, that empathic feelings are not assumed to be strictly precluded by aesthetically distanced reading. Rather, they may be productively transformed into other – i.e. more self-oriented (Cupchik et al., 1998) – types of readers' affect. More research is needed to clarify these relationships, to further validate the distinction between self-oriented and stimulus-oriented affect, and to investigate its implications for the long-term effects of literary and other reading. The possible social underpinnings of distinct aesthetic responses, as observed in our qualitative study, also remain to be investigated more closely before any generalizations can be made regarding the nexus of literary fiction and empathy.

The discipline of stylistics is ideally suited to answering these questions in more naturalistic research designs (see Gavins, 2014), unconstrained by the experimenter's practical obligation to artificially polarize texts into those possessing more or less of an isolated quality (e.g. foregrounding). Qualitative experimental research such as the present study, in turn, contributes a level of detail in pairing stimuli with verbal responses that is difficult to find elsewhere. The natural next step in bridging experimental and naturalistic approaches to reader response research is designing a community-specific qualitative experiment based on salient response patterns previously found in the field (e.g. online or face-to-face discussions). Its findings would not only enrich our knowledge of the effects of a given textual feature, but also help overcome the challenges of collecting reader response data in laboratory environments (see also Kuzmičová, 2016).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was partially enabled by a networking grant from the Joint Committee for Nordic Research Councils in the Humanities and Social Sciences (NOS-HS Grant ES521054).

Notes

1. Transportation (Green and Brock, 2000) is a psychometric construct comprising the degrees of attention, emotion, and mental imagery elicited in a narrative experience.
2. The larger experiment was carried out cross-nationally and the following researchers also contributed to the experimental design: Karin Kukkonen, University of Oslo; Lene Lauridsen, Aarhus University; Skans Kersti Nilsson, University of Borås; Torsten Pettersson, Uppsala University; Jolin Slotte, Åbo University Academy; Mette Steenberg, Aarhus University; Lisbeth Stenberg, Gothenburg University; Cecilia Therman, University of Helsinki.
3. The ART was excluded from the design because of its limited reliability in cross-national contexts.
4. Note that Cupchik et al. (1998) refrain from systematic reference to the notion of empathy proper. However, their definition of fresh emotions coincides with common definitions of reader–character empathy. For an alternative framework introducing “distance” as a key variable in emotional reader response, see Sklar (2013).

References

- Acheson DJ, Wells JB and MacDonald MC (2008) New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods* 40(1): 278–289.
- Allington D (2011) ‘It actually painted a picture of the village and the sea and the bottom of the sea’: Reading groups, cultural legitimacy, and description in narrative (with particular reference to John Steinbeck’s *The Pearl*). *Language and Literature* 20(4): 317–332.
- Bal PM and Veltkamp M (2013) How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *PLoS ONE* 8(1): e55341.
- Baron-Cohen S, Wheelwright S, Hill J, et al. (2001) The ‘Reading the Mind in the Eyes’ test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry* 42(2): 241–251.
- Belluck P (2013) For better social skills, scientists recommend a little Chekhov. *New York Times*. Available at: <http://well.blogs.nytimes.com/2013/10/03/i-know-how-youre-feeling-i-read-chekhov/> (accessed 2 February 2016).
- Bortolussi M and Dixon P (2003) *Psychonarratology: Foundations for the Empirical Study of Literary Response*. Cambridge: Cambridge University Press.
- Busselle R and Bilandzic H (2009) Measuring narrative engagement. *Media Psychology* 12(4): 321–347.
- Cupchik GC (2002) The evolution of psychical distance as an aesthetic concept. *Culture and Psychology* 8(2): 155–187.
- Cupchik GC, Oatley K and Vorderer P (1998) Emotional effects of reading excerpts from short stories by James Joyce. *Poetics* 25(6): 363–377.
- Djikic M, Oatley K and Moldoveanu MC (2013) Reading other minds: Effects of literature on empathy. *Scientific Study of Literature* 3(1): 28–47.
- Fialho O (2007) Foregrounding and readers’ refamiliarization: Understanding readers’ response to literary text. *Language and Literature* 16(2): 105–123.

- Fialho O (2012) *Self-modifying experiences in literary reading: A model for reader response*. PhD Thesis, University of Alberta, Canada. Available at: <https://era.library.ualberta.ca/public/view/item/uuid:717e43d6-57f6-4092-b783-b05ac0c4be46/> (accessed 10 October 2014).
- Fialho O, Zyngier S and Miall D (2011) Interpretation and experience: Two pedagogical interventions observed. *English in Education* 45(3): 236–253.
- Gavins J (2014) Defamiliarisation. In: Stockwell P and Whiteley S (eds) *The Cambridge Handbook of Stylistics*. Cambridge: Cambridge University Press, pp.196–211.
- Green MC and Brock TC (2000) The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology* 79(5): 701–721.
- Hakemulder F (2008) Imagining what could happen: Effects of taking the role of a character on social cognition. In: Zyngier S, Bortolussi M, Chesnokova A, et al. (eds) *Directions in Empirical Literary Studies: In Honor of Willie van Peer*. Amsterdam: John Benjamins Publishing, pp.139–153.
- Hakemulder F, Koopman E, Fialho O, et al. (2016) Learning from literature: Empirical research on readers in schools and at the workplace. In: Burke M, Fialho O and Zyngier S (eds) *Scientific Approaches to Literature in Learning Environments*. Amsterdam: John Benjamins Publishing, pp.19–38.
- Hakemulder J (2000) *The Moral Laboratory: Experiments Examining the Effects of Reading Literature on Social Perception and Moral Self-concept*. Amsterdam: John Benjamins Publishing.
- Hakemulder JF (2004) Foregrounding and its effect on readers' perception. *Discourse Processes* 38(2): 193–218.
- Johnson DR (2013) Transportation into literary fiction reduces prejudice against and increases empathy for Arab-Muslims. *Scientific Study of Literature* 3(1): 77–92.
- Kidd DC and Castano E (2013a) Reading literary fiction improves theory of mind. *Science* 342(6156): 377–380.
- Kidd DC and Castano E (2013b) Reading literary fiction improves theory of mind: Supplementary materials. *Science* 342(6156). Available at: <http://science.sciencemag.org/content/suppl/2013/10/03/science.1239918.DC1> (accessed 2 February 2016).
- Koopman EM (2015) Empathic reactions after reading: The role of genre, personal factors and affective responses. *Poetics* 50(6): 62–79.
- Kuijpers MM, Hakemulder F, Tan ES, et al. (2014) Exploring absorbing reading experiences: Developing and validating a self-report scale to measure story world absorption. *Scientific Study of Literature* 4(1): 89–122.
- Kuivalainen P (2009) Emotions in narrative: A linguistic study of Katherine Mansfield's short fiction. *Helsinki English Studies: Electronic Journal* 5. Available at: <http://blogs.helsinki.fi/hes-eng/volumes/volume-5/emotions-in-narrative-a-linguistic-study-of-katherine-mansfield%E2%80%99s-short-fiction-paivi-kuivalainen/> (accessed 3 February 2016).
- Kuzmičová A (2012) Presence in the reading of literary narrative: A case for motor enactment. *Semiotica* 189(1): 23–48.
- Kuzmičová A (2016) Does it matter where you read? Situating narrative in physical environment. *Communication Theory* 26(3): 290–308.
- Mansfield K (1950) Fluen. In: *Havefesten: utvalgte noveller*. Oslo: Gyldendal, pp.253–258.
- Mansfield K (1923) The Fly. In: *The Dove's Nest and Other Stories*. London: Constable and Co., pp.45–54.
- Mar RA, Oatley K, Hirsh J, et al. (2006) Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality* 40(5): 694–712.

- Mar RA, Oatley K and Peterson JB (2009) Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communications* 34(4): 407–428.
- Miall DS (2006) *Literary Reading: Empirical and Theoretical Studies*. New York: Peter Lang.
- Miall DS (2011) Emotions and the structuring of narrative responses. *Poetics Today* 32(2): 324–348.
- Miall DS and Kuiken D (1994) Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics* 22(5): 389–407.
- Miall DS and Kuiken D (1995) Aspects of literary response: A new questionnaire. *Research in the Teaching of English* 29(1): 37–58.
- OECD (2002) Reading unit 8: The gift. In: *Sample Tasks from the PISA 2000 Assessment: Reading, Mathematical and Scientific Literacy*. Paris: OECD Publishing, pp.60–62.
- van Peer W (1986) *Stylistics and Psychology: Investigations of Foregrounding*. London: Taylor and Francis.
- Rosch E (1978) Principles of categorization. In: Rosch E and Lloyd BB (eds) *Cognition and Categorization*. Hillsdale: Lawrence Erlbaum, pp.27–48.
- Shamay-Tsoory SG and Aharon-Peretz J (2007) Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia* 45(13): 3054–3067.
- Shklovsky V (1988/1925) Art as technique. In: Lodge D (ed.), Lemon LT and Reis MJ (trans) *Modern Criticism and Theory: A Reader*. London: Longman, pp.16–30.
- Sikora S, Kuiken D and Miall DS (2011) Expressive reading: A phenomenological study of readers' experience of Coleridge's *The Rime of the Ancient Mariner*. *Psychology of Aesthetics, Creativity, and the Arts* 5(3): 258–268.
- Sklar H (2013) *The Art of Sympathy in Fiction: Forms of Ethical and Emotional Persuasion*. Amsterdam: John Benjamins Publishing.
- Sommerfeldt B and Skårderud F (2008) Adult Reading the Mind in the Eyes test: Norwegian version. Available at: www.autismresearchcentre.com/arc_tests/ (accessed 2 February 2016).
- Stanovich KE and West RF (1989) Exposure to print and orthographic processing. *Reading Research Quarterly* 24(4): 402–433.
- Torkildsen T (2014) Fluen: Unpublished adaptation.

Author biographies

Anežka Kuzmičová is a postdoctoral research fellow in the Department of Culture and Aesthetics, Stockholm University, Sweden. She has published on a wide range of topics relating to the act of reading, e.g. readers' mental imagery, digital audiobooks and the physical reading environment.

Anne Mangen is a professor of literacy at The Reading Centre, University of Stavanger, Norway. Her research interests include the effects of digitization on cognitive and emotional aspects of literary reading.

Hildegunn Støle is a former teacher and holds a PhD in English historical pragmatics. Employed at The Reading Centre, University of Stavanger, she works with children's reading, education, and reading assessment.

Anne Charlotte Begnum is an associate professor of special education at The Reading Centre, University of Stavanger.