ORIGINAL RESEARCH REPORT

The Influence of Mental Imagery Instructions and Personality Characteristics on Reading Experiences

Marloes Mak*, Clarissa de Vries* and Roel M. Willems*,†,‡

It is well established that readers form mental images when reading a narrative. However, the consequences of mental imagery (i.e. the influence of mental imagery on the way people experience stories) are still unclear. Here we manipulated the amount of mental imagery that participants engaged in while reading short literary stories in two experiments. Participants received pre-reading instructions aimed at encouraging or discouraging mental imagery. After reading, participants answered questions about their reading experiences. We also measured individual trait differences that are relevant for literary reading experiences. The results from the first experiment suggests an important role of mental imagery in determining reading experiences. However, the results from the second experiment show that mental imagery is only a weak predictor of reading experiences compared to individual (trait) differences in how imaginative participants were. Moreover, the influence of mental imagery instructions did not extend to reading experiences unrelated to mental imagery. The implications of these results for the relationship between mental imagery and reading experiences are discussed.

Introduction
It is well established that readers perceive mental images during reading (Green & Brock, 2000; Jacobs, 2015). For instance, an eye tracking study showed that people are responsive to mental simulation-eliciting content in stories (Mak & Willems, 2019). It was found that when participants were reading action descriptions (assumed to elicit motor simulation) they sped up, whereas they slowed down when reading perceptual descriptions or mental event descriptions (assumed to elicit perceptual simulation or mentalizing, respectively).

Additionally, there is a relationship between the amount of imagery and subjective experiences during reading. Mak and Willems (2019) found that individual differences in the responsiveness to simulation-eliciting content were related to participants' subjective experiences (such as absorption and appreciation). This is only one example of work showing that mental simulation during reading is associated with absorption in and appreciation of stories (see also Green, 2004; Green & Brock, 2002; Kuijpers, Hakemulder, Tan, & Doicaru, 2014; Mol & Jolles, 2014; Weibel, Wissmath, & Mast, 2011).

Next to individual variation in amount of imagery perceived during reading, there are a number of stable (personality) characteristics that are associated with reading experiences. In the experiments we report here, we decided to study both the role of instructed mental imagery and the role of individual (trait) differences in literary reading. Below, we will discuss the relationships between (1) mental imagery and reading experiences and between (2) individual (trait) differences and reading experiences, before we (3) introduce the set-up and hypotheses of our experiment.

Mental Imagery and Reading Experiences
As mentioned above, people engage in mental imagery[1] when reading stories, and mental imagery is an important driver of absorption: it has been found that visualizing the story world will strengthen people's experience of absorption in a story (Green & Brock, 2002; Kuijpers et al., 2014; Kuiken & Douglas, 2017).[2] Absorption has been defined by Kuijpers and colleagues (2014) as "the *subjective experience* of being absorbed in the story world of a *narrative text*" (p. 90, emphasis in the original). It describes the feeling we may have when reading a good story or book where we go beyond comprehending the meaning of the words on a page, to a captivating experience that can help us become completely involved in the stories we read. This experience has been reported on in widely diverging disciplines, and has also been defined as transportation (Gerrig, 1993; Green & Brock, 2000), narrative engagement (Busselle & Bilandzic, 2008),

* Centre for Language Studies, Radboud University Nijmegen, Erasmusplein, Nijmegen, NL

† Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Kapittelweg, Nijmegen, NL

‡ Max Planck Institute for Psycholinguistics, Nijmegen, Wundtlaan, Nijmegen, NL

Corresponding author: Marloes Mak (H.Mak@let.ru.nl)

narrative presence (Kuzmičová, 2012), immersion (Ryan, 2001; see also Jacobs, 2015) or flow (Csikszentmihalyi, 1990). For the sake of clarity, we will refer to this experience as absorption for the remainder of this article.

Absorption is a multifaceted construct that tries to capture the entirety of the subjective experience of being captivated by a good story (Kuijpers et al., 2014). It is proposed that absorption consists of multiple dimensions, being mental imagery, emotional engagement, attention, and transportation (Kuijpers et al., 2014). Mental imagery (as a dimension of absorption) is defined as a visualization of the *story world*, whereas emotional engagement could be seen as the emotional counterpart of mental imagery: the sympathetic and empathic feelings for the *characters* in the story. Attention is characterized as a heightened focus or concentration of the reader towards the story world – and as a consequence a lower concentration towards the here and now. Transportation is seen as the feeling a reader can have of being part of the story world as opposed to the real world. Together these 4 dimensions can result in an experience of complete absorption in a narrative or story world.[3]

Another connection between absorption and mental imagery, is that they have both been found to be associated with another important aspect of people's reading experiences: the enjoyment (or appreciation) of stories (Busselle & Bilandzic, 2009; Green, 2004; Green, Brock & Kaufman, 2004; Kuijpers et al., 2014; Kuiken & Douglas, 2017; Mol & Jolles, 2014; Weibel et al., 2011). As there does not seem to be strong consensus with regard to the definition of appreciation (especially between different disciplines, e.g., communication research and literature studies), in the current experiment we have considered both the overall enjoyment of narratives and other facets of aesthetic experiences that we believe could play a role in our enjoyment and appreciation of stories (e.g., whether a reader is emotionally moved by a story, or finds it amusing). To test this multitude of facets of aesthetic experiences, we used adjectives that are often used to describe the multiple dimensions of aesthetic experiences (Knoop, Wagner, Jacobsen, & Menninghaus, 2016). These adjectives were taken from a list of adjectives that were most often used by people to describe their experience while reading literature (Knoop et al., 2016). Because this list of adjectives was compiled in a "bottom-up" fashion (i.e. the adjectives are derived directly from the experiences of readers), it was assumed that these adjectives could successfully tap into multiple facets of aesthetic experiences of the readers in our experiments. We found it important to look at appreciation apart from absorption, because we wanted to consider a measure of reading experience that was more distantly related (but not unrelated) to mental imagery than absorption (recall that mental imagery is considered to be one of the four subcomponents of absorption).

Reading instructions and mental imagery
In order to test the relationship between mental imagery and reading experiences, we wanted to make sure that some readers in our experiments would engage in mental imagery more than others. To this effect, we employed a method from a related but separate line of research, where it was found that instructing students to create images of what they had read impacts text comprehension (see De Koning & van der Schoot, 2013 for an extensive overview). Apart from text comprehension, some studies have also found direct links between pre-reading instructions and reading experiences. Green and Brock (2000) found that instructing readers to judge the difficulty of a story to establish the suitability for fourth-grade readers, led to lower experienced transportation (which is conceptually comparable to absorption) in comparison to readers who were instructed to pay specific attention to the story plot. The rationale of the study was that instructing participants to pay attention to the suitability of the text for fourth-graders, would lead to less absorption than instructing them to focus on the story plot. In a follow-up experiment, Green (2004) subsequently instructed participants to use relaxation strategies during reading to increase transportation. This manipulation did however not lead more experienced transportation in these participants, when compared to participants who had received a neutral instruction. Perhaps this instruction was not associated with higher transportation because relaxation is relatively unrelated to the process of transportation. Therefore, the advice for future research was to use a pre-reading instruction focusing on a specific component of transportation (e.g., imagery; Green, 2004, p. 261). Johnson, Cushman, Borden and McCune (2013) made a successful first attempt in this direction. Instead of pre-reading instructions, they gave participants an imagery generation training, which subsequently resulted in increased transportation when reading a narrative. In the current study, we take these findings as the starting point for our investigation of the influence of explicit imagery instructions on subjective reading experiences.

Participants in the experiments described in the current paper were either encouraged or discouraged to engage in mental imagery through pre-reading instructions, after which their subjective experiences while reading literary short stories were measured. Even though pre-reading instructions have not been successful in all studies attempting to influence reading experiences, Tukachinsky (2014) noted in a review of these studies that the effect of pre-reading instructions seemed quite reliable. Moreover, we used an instruction specifically targeting mental imagery, which was suggested by Green (2004) as possibly more powerful in manipulating reading experiences than instructions aimed at more general processes. The purpose of this specific mental imagery instruction was to manipulate the amount of mental imagery between participants in order to establish what role mental imagery plays in subjective reading experiences.

Individual (Trait) Differences in Reading Experiences
As mentioned above, mental imagery is not the only factor that plays a role in reading experiences. Variation in experienced absorption and appreciation can also be due to individual differences in situational factors (e.g., stress, mood, distractions, level of energy), or more stable

characteristics. For example, amount of print exposure is negatively associated with reading difficulties (Stanovich & West, 1989), positively associated with reading skills (Acheson, Wells, & MacDonald, 2008) and positively associated with language ability, school success, Theory of Mind and empathy (Brysbaert, Sui, Dirix, & Hintz, 2020). Additionally, reading habits in daily life are closely related to reading experiences: more habitual readers experience more absorption (Kuijpers, Douglas, & Kuiken, 2018) and enjoy reading more (Mol & Jolles, 2014).

Furthermore, a range of personality characteristics have been found to be associated with reading experiences. Individuals reporting more need for affect (Appel & Richter, 2010), as well as more transportable individuals (Bilandzic & Busselle, 2011), reported experiencing more transportation while reading a story or watching a film. Similarly, Need for Cognition was found to be a predictor of transportation experienced while reading stories (Green et al., 2008) or watching films (Hall & Zwarun, 2012). Openness was positively associated with reported interest for stories (which is related to enjoyment/appreciation of stories; Fayn, Tiliopoulos, & MacCann, 2015), absorption (although indirectly, via reading habits; Kuijpers et al., 2018), and the overall likelihood that people read literature for leisure (Kraaykamp & Eijck, 2005; see also Schutte & Malouff, 2004). Interestingly, Malanchini and colleagues (2017) link differences in reading enjoyment and motivation to genetic differences.

Because of the important role of the abovementioned individual (trait) differences in reading experiences, we took these into account in the experiments reported here. Because we controlled for the role of these individual (trait) differences in reading experiences when studying the role of guided mental imagery instructions in reading experiences, we will be able to draw conclusions about the role of mental imagery instructions *over and above* individual (trait) differences from the results of our experiments.

Current Study
In the current study we investigate the respective roles of mental imagery and individual differences on subjective experiences during reading. We test in two experimental studies whether guided mental imagery instructions influenced reading experiences. In keeping with the suggestion made by Green (2004), we tested the effect of pre-reading instructions specifically focusing on mental imagery in our experiments.

In our first experiment participants read a literary short story and subsequently rated their reading experiences on several questionnaires. One group of participants was instructed to use mental imagery while reading the stories, whereas another group of participants was asked to read the stories for leisure. In the second experiment a third instruction was added, designed to distract participants from the plot of the story. This third instruction was added to control for a task confound in the first experiment, where the imagery instruction was more effortful to follow than the leisure instruction. The task in the third instruction was as effortful to complete

as the task in the imagery instruction, but it was designed to distract participants from the plot of the story, and therefore to make their mental imagery less vivid. With this experiment we wanted to test whether overall reading experiences can be modified using reading instructions focusing on one specific facet of these reading experiences and whether these reading instructions can "overrule" the influence of individual (trait) differences (e.g., reading habits, personality characteristics) on reading experiences. Based on the literature we hypothesized that mental imagery instructions would result in more mental imagery compared to the control group (leisure readers) and therefore in more absorption and appreciation. In contrast, we hypothesized that the distracting instruction added in the second experiment would result in less mental imagery and therefore in reduced absorption and appreciation.

If specific reading instructions would indeed prove powerful in altering reading experiences these could be used to promote reading in people who do not read for leisure, which could have positive consequences for among others school success (Chiu & McBride-Chang, 2006; Mol & Jolles, 2014; Retelsdorf, Köller, & Möller, 2011) second language learning (Lao & Krashen, 2000; Lee, Schallert, & Kim, 2015; Yamashita, 2008), social cognition and empathy (e.g., Fong, Mullin, & Mar, 2013; Johnson et al., 2013; Mar & Oatley, 2008; Oatley, 2016), or persuasion (e.g., Dal Cin, Zanna, & Fong, 2004; Green & Brock, 2000). If, however, other factors (e.g., individual trait differences) are found to be a stronger driver of absorption and appreciation than imagery instructions, this would indicate that using such instructions in for instance educational settings is not an optimal intervention to increase reading pleasure. A third option would be that individual trait differences and imagery instructions interact as drivers of absorption and appreciation. In that case, this could indicate that using imagery instructions would only be useful for some individuals and that it would be necessary to find a way to determine which individuals would or would not benefit from such instructions.

Experiment 1
Methods
This first experiment was conducted in the context of Bachelor's theses, for which five students of Communication and Information Studies at the Radboud University in Nijmegen, The Netherlands, worked together under the supervision of the first author to conduct an experiment testing the influence of mental imagery-inducing reading instructions on reading experiences.

Participants
A total of 120 participants took part in this first experiment. To ensure that participants understood their instructions for the experiment they were asked to repeat what they had been instructed to do while reading (i.e. mental imagery versus reading for leisure) after reading the story. Due to an error during data collection, the experimental condition of 20 participants was not registered correctly. Data from these 20 participants

were excluded from analysis. It was double checked that the data from the remaining 100 participants were registered correctly before moving on with the analyses. The remaining participants consisted of an experimental group (n = 45; 25 females; Age: $M(SD)$ = 32 (15) years old; age range = 19–71) and a control group (n = 55; 33 females; Age: $M(SD)$ = 33 (16) years old; age range = 17–82). Chi-square tests indicated that there were no significant differences between the participants in the two groups with respect to biological gender ($\chi^2(1)$ = 0.06; p = .81) and educational level ($\chi^2(1)$ = 10.06; p = .07). However, to control for any possible individual differences in age, biological gender or educational level (measured as the highest completed education, ranging from 1 to 6, where 6 was the highest possible level of education in the Netherlands), we considered these factors as covariates in our analyses.

Prior to the experiment, participants were informed about the procedure of the experiment. It was made clear that participation was voluntary and that it was allowed to withdraw from the experiment at any time without need for explanation. All participants gave written informed consent in accordance with the Declaration of Helsinki. The study was approved by the local ethics committee.

Materials

Materials used in Experiment 1 consisted of the story that was read, the instructions participants received before reading, and the questionnaires participants filled in after reading. We will now discuss these materials in more detail.

Story

The story used in the first experiment was an existing literary short story by the acclaimed Dutch writer Rob van Essen (2014), called *De mensen die alles lieten bezorgen* (The people who had everything delivered). The story was 2988 words long and took participants about 10–15 minutes to read. The story recounts the experiences of a man who lives in an apartment building in Amsterdam. His neighbors rent out their apartment while they are on holiday for the Christmas days, and a morbidly obese British couple stays there. When the wife has a heart attack she has to be lifted out of the apartment by a firetruck, as there is no elevator in the building. The events in the story are described using very colorful language and are easy to visualize.

Instructions

Before reading the stories, participants received a reading instruction. The experimental group was instructed to *"Pay close attention to the story and read the story as you would normally read a story. Use your imagination while reading, by visualizing the surroundings described in the story and envisioning the actions of the characters. Imagine the main character standing in front of you, imagine what happens and pay close attention to what all of the characters are doing"*. The control group was simply instructed to *"Pay close attention to the story and read the story as you would normally read a story"*. This way both groups were instructed to read the story attentively, but the experimental group was encouraged to form a vivid mental image of the events described in the story. When participants had finished reading, they were asked to repeat their reading instruction to the experimenter. If they were unable to repeat the instruction correctly or if their answer indicated they did not use the instruction while reading the story, data for these participants were excluded from the analysis.

Questionnaires

After reading, participants had to fill in some questionnaires measuring their reading experiences (i.e. Story World Absorption, see Kuijpers et al., 2014), reading habits in daily life, and other more general information (i.e. age, level of education, and biological gender). Since this experiment was conducted in the context of Bachelor's theses, a couple of questionnaires were devised by the Communication and Information Studies students regarding topics not under study here (e.g., attitudes towards fast food, behavioral intentions regarding healthy eating). Since we do not have specific theoretical assumptions regarding these topics, we will not discuss these measures in the current paper.

Story world absorption was measured using the Story World Absorption Scale (SWAS; Kuijpers et al., 2014). The SWAS is a validated scale consisting of 18 items with high internal validity (Kuijpers et al., 2014) which measures 4 aspects of story world absorption on the four subscales Attention, Transportation, Emotional Engagement and Mental Imagery (e.g., *When I finished the story I was surprised to see that time had gone by so fast; I could imagine what the world in which the story took place looked like*). Participants rated each question on a 7-point scale (1 = disagree, 7 = agree).

Reading habits were measured using five multiple choice questions about reading habits in everyday life, with 4 or 5 optional answers (Hartung, Burke, Hagoort, & Willems, 2016; Mak & Willems, 2019; e.g., *How often do you read fiction*; *How many books do you read each year*). Additionally, participants were asked for their genre preference in an open-ended question, where they could list up to three preferred genres.

Procedure

Participants were recruited and tested in a quiet room at the university campus or at home. Participants were informed about the procedure and were asked for written informed consent. At the start of the experiment, participants were given one of the two possible instructions on paper (as described above). If necessary, the instruction was clarified by the experimenter. After that, participants read the story at their own pace. Both the instructions and the story were read from paper. After reading, participants were asked to fill in the questionnaires (SWAS, reading habits, general information) on the experimenter's laptop. The entire procedure took about 20–25 minutes.

Data-analysis

Data-analysis was done using the '*stats*' package in R version 3.6.1 (R Core Team, 2018). We constructed a linear regression model that predicted average scores on the

SWAS based on experimental group (imagery instruction contrasted with control). Biological gender (male contrasted with female), age, level of education, and reading habits were added to the model as general variables expected to explain additional variance. As a result, any effects of the experimental group would represent variance explained by the given instruction over and above variance explained by these demographic variables. Similar models were constructed to predict scores on the four subscales of the SWAS separately (i.e. Attention, Transportation, Emotional Engagement and Mental Imagery). All continuous predictors were centered and scaled. Variance Inflation Factors (VIFs) were calculated for all models, to check for multicollinearity between predictors. All VIFs for all models were between 1 and 2, indicating that multicollinearity was not problematic in our models and all planned predictors could be entered into the models.

Results

Questionnaires

To test the reliability of the used scales and subscales, Ω_t was calculated. We decided to calculate Ω as opposed to Cronbach's α, since it has been argued by several researchers that Ω is a more appropriate measure of reliability (e.g., Dunn, Baguley, & Brunsden, 2014; Peters, 2014; Revelle, 2014). We decided to report Ω_t as opposed to Ω_h because we assumed the constructs measured with our scale to be multidimensional (see Revelle, 2014). The average scores on the Story World Absorption Scale as well as the scores on the four subscales of the SWAS

all showed sufficient to excellent reliability; total SWAS (18 items), Ω_t = .95; SWAS Attention (5 items), Ω_t = .87; Transportation (5 items), Ω_t = .84; Emotional Engagement (5 items), Ω_t = .92; Mental Imagery (3 items), Ω_t = .79. Descriptive statistics for the questionnaire scores per subscale and per group are visualized in **Figure 1**.

The answers on the reading habits questionnaire were measured on a scale ranging from 1 to 5 on four of the five multiple choice questions, but from 1 to 4 on the final question. Therefore, z-scores were calculated for all questions on this questionnaire (with higher scores for more habitual readers). Overall reliability was sufficient, Ω_t = .78.

Main analysis

The model predicting average scores on the SWAS based on biological gender, age, level of education, reading habits and experimental group (imagery instruction versus control; model adjusted R^2 = 0.185) showed that participants in the experimental group were more absorbed than participants in the control group (see **Table 1.1** for all results of this model; see **Figure 1A**). Additionally, participants who were more habitual readers, also reported more story world absorption. Finally, males reported less story world absorption than females and participants with a higher level of education were more absorbed.

To find out which aspects of story world absorption this difference between the experimental and control group stems from, similar models were built to predict
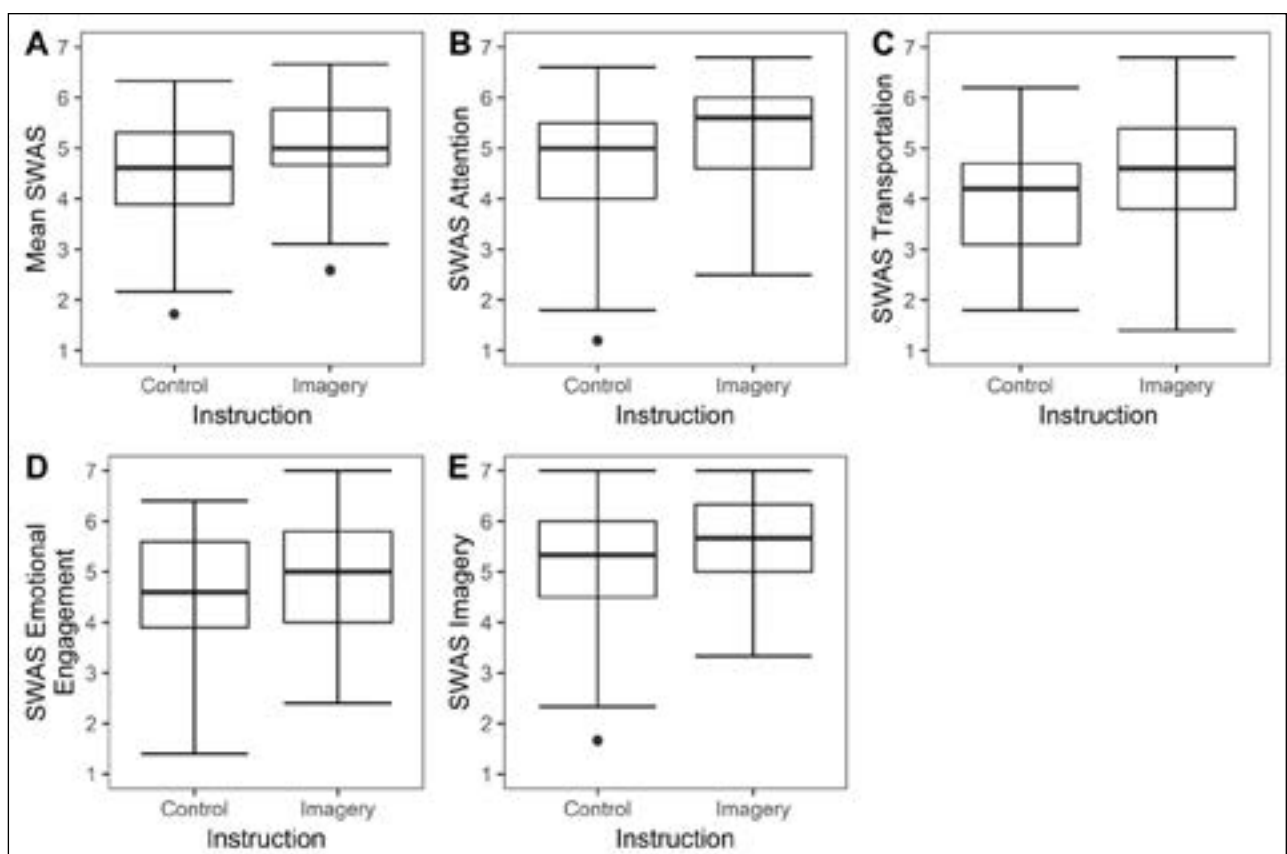
Figure 1: Boxplots depicting scores on the SWAS, per subscale and per group.

Table 1: Coefficients of the models predicting absorption based on type of instruction (mental imagery instruction contrasted with the control instruction), taking into account biological gender (male contrasted with female), age, self-reported reading habits, and level of education. Significant predictors are marked (* $p < .05$, ** $p < .01$, *** $p < .001$).

		B	*SE*	*t*-value	*p*-value
1. SWAS	(Intercept)	4.65	0.14	33.11	<.001***
	Imagery instruction	0.58	0.18	3.18	.002**
	Reading habits	0.32	.14	2.24	.03*
	Age	0.06	0.09	0.70	.49
	Biological gender (male)	−0.43	0.18	−2.39	.02*
	Level of education	0.21	0.09	2.30	.02*
1a. SWAS Attention	(Intercept)	4.86	0.16	29.45	<.001***
	Imagery instruction	0.62	0.21	2.87	.005**
	Reading habits	0.47	0.17	2.83	.006**
	Age	−0.01	0.11	−0.05	.96
	Biological gender (male)	−0.44	0.21	−2.05	.04*
	Level of education	0.20	0.11	1.83	.07
1b. SWAS Transportation	(Intercept)	3.99	0.17	23.90	<.001***
	Imagery instruction	0.62	0.22	2.84	.005**
	Reading habits	0.21	0.17	1.23	.22
	Age	0.02	0.11	0.17	.87
	Biological gender (male)	−0.31	0.22	−1.45	.15
	Level of education	0.43	0.11	3.95	<.001***
1c. SWAS Emotional Engagement	(Intercept)	4.77	0.18	26.25	<.001***
	Imagery instruction	0.45	0.24	1.90	.06
	Reading habits	0.38	0.18	2.10	.04*
	Age	0.15	0.12	1.24	.22
	Biological gender (male)	−0.58	0.23	−2.46	.02*
	Level of education	0.09	0.12	0.80	.43
1d. SWAS Mental Imagery	(Intercept)	5.22	0.17	30.01	<.001***
	Imagery instruction	0.68	0.23	3.01	.003**
	Reading habits	0.13	0.17	0.76	.45
	Age	0.11	0.11	1.02	.31
	Biological gender (male)	−0.39	0.23	−1.72	.09
	Level of education	0.06	0.11	0.54	.59

the scores on the four subscales of the SWAS (Attention, Transportation, Emotional Engagement, Mental Imagery). The model predicting the scores on the Attention subscale of the SWAS (model adjusted R^2 = 0.172; see **Table 1.1a**), showed that participants in the experimental group read the story more attentively (see **Figure 1B**). More habitual readers in daily life also reported more Attention to the story and males reported less Attention than females.

For the Transportation subscale of the SWAS (model adjusted R^2 = 0.170; see **Table 1.1b**), it was found that participants in the experimental group experienced more transportation than participants in the control group (see **Figure 1C**). Participants with a higher level of education also reported experiencing more transportation into the story.

The model predicting scores on the Emotional Engagement subscale of the SWAS (model adjusted R^2 = 0.118; see **Table 1.1c**) revealed no differences between the experimental and control group with regard to emotional engagement (see **Figure 1D**). Males were less emotionally engaged than females. Additionally, more habitual readers reported to be more emotionally engaged with the story.

In support of the effectiveness of our manipulation, participants in the experimental group reported more

mental imagery than participants in the control group (see **Table 1.1d**; see **Figure 1E**) on the Mental Imagery subscale of the SWAS (model adjusted $R^2 = 0.084$). None of the other tested predictors (biological gender, age, level of education, reading habits) were significantly related to scores on the Mental Imagery subscale of the SWAS.

Discussion

The results from this first experiment show that mental imagery-inducing reading instructions were associated with a stronger absorption experience, in particular a stronger attention towards the story, a stronger experience of transportation into the story world and more reported use of mental imagery (confirming the effectiveness of our manipulation). This suggests that, indeed, pre-reading instructions focusing on specific aspects of reading (such as mental imagery) can influence the way readers experience stories.

Apart from the influence of reading instructions, we found substantial individual differences in reading experiences. For instance, females reported to be more absorbed by the story. When looking at the subcomponents of absorption, it became clear that this difference between males and females was most prominent in the emotional engagement component of absorption: women were more emotionally engaged when reading the story than men. Level of education also appeared to explain some of the variation in absorption. This was mostly visible in the transportation subcomponent of absorption: participants with a higher level of education, reported experiencing more transportation into the story world. Habitual readers were also more absorbed than participants who did not read much in their own time. This was visible in their scores on the attention and emotional engagement subcomponents of absorption: more habitual readers reported more attention to the story and more emotional engagement with the story. Age was not significantly related to the absorption experience, nor to any of its subcomponents, as has also been found in previous work (Hartung, Withers, Hagoort, & Willems, 2017).

From this first experiment it could be concluded that there are indeed individual differences in reading experiences, which are related to both mental imagery-inducing reading instructions and stable individual differences (i.e. biological gender, level of education, reading habits). However, a few questions remain after this experiment. First and foremost, a difference in reading experiences between the experimental group and the control group could mean two things. Although it is very well possible that absorption was enhanced through induced mental imagery as a result of the mental imagery reading instruction in the experimental group, it could also be the case that elaborate reading instructions in general promote more intensive reading and as a result a stronger experience of absorption (in both reading instructions used in this experiment, participants were told to read the story attentively, but only in the imagery instruction participants were asked to vividly imagine the events happening in the story on top of reading the story attentively). That is, in this experiment, the imagery group got more elaborate (longer) instructions than the control group and this could have led to the differences that we observed (see De Koning & van der Schoot, 2013 for a similar argument). Secondly, the reading instruction was aimed at mental imagery, which is part of the absorption experience. Although the effect of the reading instruction did translate to other subcomponents of absorption (attention, transportation), it would be interesting to find out if a mental imagery reading instruction could also influence other reading experiences, most notably story appreciation. Third, we did look at some general individual differences that could influence reading experiences, such as biological gender, age and reading experience, but we did not look at any personality characteristics that have been associated with reading experiences before. In the next experiment we have therefore considered the personality characteristics fantasy (how imaginative a person is) and perspective taking (the extent to which someone takes other peoples' perspectives in daily life).

To be able to answer these remaining questions, we ran a second experiment, where we made the mental imagery instruction more clear (to improve compliance with the instruction) and included a third instruction, which was as elaborate as the mental imagery instruction, but its content was aimed at diverting the reader from the narrative's plot (but still encouraging thorough reading of the text; adapted from an instruction successfully used to lower experienced transportation; Green & Brock, 2000). This way, we could rule out that the effect of instruction was simply the result of more intensive reading instead of being related to the actual content of the instruction. To elaborate, we added this third instruction to test the alternative explanation for the results from experiment 1: that elaborate, detailed reading instructions in general promote more intensive reading and as a result a stronger experience of absorption. If both detailed instructions would increase absorption, this would be evidence that this increase in absorption is the result of more thorough reading due to the length or details in the instructions. However, if only the imagery instruction would increase absorption and the other detailed instruction would not change or even decrease absorption, we would have stronger evidence for our hypothesis that the increase in absorption after the imagery instruction is the result of the *content* of the instruction. For this third instruction (that was as detailed as the imagery instruction), we chose an instruction aimed at *decreasing* absorption since using a second detailed instruction that was aimed at *increasing* absorption would not have been helpful: if in this case both detailed instructions would increase absorption, we would still not know whether this was due to the content of the instructions, and not simply to the fact that they were detailed instructions that encouraged thorough reading.

We also included measures for story appreciation, a more thorough measure of mental imagery, and measures of personality traits we thought might play a role in reading experiences. Furthermore, to ensure more experimental control we tested our experiment in a more controlled environment in this second study. Finally, we decided to

use two new stories to find out if the effect of reading instructions would also extend to different stories.

Experiment 2
Methods
In this experiment participants were divided into three groups. Apart from the group receiving an elaborate mental imagery instruction and the control group, we also included a group that received the instruction to judge whether the writing style of the story (sentence construction, word use) was suitable for teenagers of about 14 or 15 years old, who were in the lower grades of Dutch secondary education (henceforth called secondary school suitability instruction; cf. Green & Brock, 2000). Such "distraction manipulations" have in previous studies been particularly useful in manipulating transportation (for a review, see Tukachinsky, 2014). If the *length* of the imagery instruction was the reason people became more transported in the first experiment, this secondary school suitability instruction should also result in increased absorption and there should be no difference between the results for the imagery instruction and the suitability instruction. However, if the effect of the instruction was due to the *content* of the instruction, the imagery instruction should increase absorption but the suitability instruction should decrease absorption. This enabled us to test if an effect of instruction was the result of the actual content of the instruction or, alternatively, was simply the result of more intensive reading.

Participants
To ensure sufficient statistical power the sample size of this study was calculated in G-power (Faul, Erdfelder, Buchner, & Lang, 2009) using the effect size from experiment 1. This resulted in a required sample size of approximately 120 for an estimated power of .85, divided over 3 groups. A total of 125 participants (102 females) participated in the second experiment. The data of 7 participants had to be discarded because of procedural errors (4), too much missing data (2), or because they did not have enough time to finish the experiment. The remaining 118 participants (99 females) were split up in a group receiving a mental imagery instruction (n = 39; 32 females; M age = 24 years old), a group receiving a secondary school suitability instruction (n = 39; 35 females; M age = 23 years old) and a control group (n = 40; 32 females; M age = 24 years old). There were no differences between the participants in the three groups with respect to biological gender ($\chi^2(2)$ = 3.07; p = .22), nor were there differences between groups in age ($F(2, 233)$ = 0.77, p = .46), reading habits (self-report: $F(2, 233)$ = 0.17, p = .84; ART-score: $F(2, 233)$ = .21, p = .81) or personality characteristics (IRI Fantasy: $F(2, 233)$ = 2.81, p = .06; IRI Perspective Taking: $F(2, 233)$ = 0.43, p = .65; see below for an extensive description of all used questionnaires). Participants were all healthy, native speakers of Dutch, without dyslexia.

Participants were recruited from the participant database of the Radboud University and received €10 or course credit for participation in this study. Prior to the experiment, participants were informed about the procedure of the experiment. It was made clear that participation was voluntary and that it was allowed to withdraw from the experiment at any time without need for explanation. All participants gave written informed consent in accordance with the Declaration of Helsinki and the study was approved by the local ethics committee.

Materials
Materials used in Experiment 2 consisted of the two stories that were read, the instructions participants received before reading, and the questionnaires participants filled in after reading. We will now discuss these materials in more detail.

Stories
Instead of just one story, participants read two stories in the second experiment. Both were literary short stories written by acclaimed Dutch writers. The first story, *Brommer op zee* (Moped on sea), was written by Maarten Biesheuvel (1972) and was 1827 words long. It is a surrealistic story about a boy on a boat and his encounter with a man riding a moped at sea in the middle of the night. The second story, *God en de gekkenrechter* (God and the judge of the insane), was written by Adriaan van Dis (1986) and was 2026 words long. In this story, the author narrates the story of a mentally instable man who is convinced that he is God, and believes that therefore all his excrements are holy and should not be thrown away. Apart from that, he terrorizes the neighborhood, leading to his institutionalization later on in the story, after which he finally seems to realize that he was mistaken in thinking that he was God. Both stories contain many descriptions that could guide mental imagery of the stories. It took participants about 15 minutes to read each story. In the remainder of the section about Experiment 2 the story *Moped on sea* will be referred to as Story A, and *God and the judge of the insane* will be referred to as Story B. Note that the stories were read in counterbalanced order: Half of the participants started with Story A and half of the participants started with Story B.

Instructions
Before reading the stories, participants were given a reading instruction. Every participant received either a mental imagery instruction, a secondary school suitability instruction, or a neutral control instruction. After participants had read the first story, they received a short reminder of their reading instruction to urge them to also keep the instruction in mind while reading the second story.

The group receiving the mental imagery instruction was told *"In a short while, you will be reading a short story. During reading, try to vividly imagine the events happening in the story. Vividly imagine what you see, hear, feel or smell. For example, envision the characters and places described in the story, imagine what the conversations and environmental sounds sound like, what the odors smell like, how the physical experiences of the characters feel".*

The group receiving the secondary school suitability instruction was told *"In a short while, you will be reading a short story. Your job is to make sure the text is suitable for*

students in the lower grades of secondary school, of about 14 or 15 years old. The content of the story is not important, please pay attention to the writing style: the sentence constructions and the word use of the author of the story. Try to focus on these two aspects while reading the story. Determine whether the word use and sentence constructions are of a suitable level for students in the lower grades of secondary school".

The control group, who received a short, neutral instruction, was told *"In a short while, you will be reading a short story. Please read this story the way you would usually read a story for your leisure".*

To make sure all participants understood their reading instruction, they were presented with a short excerpt from a different story (which was stylistically comparable to the two experimental stories). Participants had to apply the reading instruction while reading this fragment and were afterwards asked to check with themselves whether they indeed applied the reading instructions while reading. After participants had practiced the instruction on the example fragment the instructions were repeated and participants were told to start reading the stories. Participants were reminded that they were allowed to read at their own pace and did not have to hurry. As mentioned previously, participants were given another reminder of the reading instruction just before they started reading the second story. The motivation behind this reminder was that participants had to fill in some questionnaires

about their reading experience after reading the first story. Therefore, we wanted to make sure they would still remember the instruction while reading the second story. Although participants received different instructions, they all practiced their instruction on the same fragment and received reminders of their instruction at the same moments. This way, we tried to make sure that all three groups would read the story equally attentively.

Questionnaires
Just as in experiment 1, we asked participants to fill in questionnaires (discussed in detail below) regarding their reading experience after reading each story. However, this time we did not only use the Story World Absorption Scale to measure reading experience, but also a questionnaire measuring story appreciation and a questionnaire measuring the vividness of the imagery experienced during reading (for an overview of all questionnaires used in this experiment, see **Figure 2**). After reading both stories and filling in the story-related questionnaires, participants were asked to fill in some additional questionnaires measuring reading habits in daily life, personality characteristics, story comprehension and more general information (i.e. age and biological gender). Level of education was not considered in this experiment as nearly all of the participants were university students and therefore no claims could be made about the role of level of education in reading experiences from this experiment. We will discuss the questionnaires
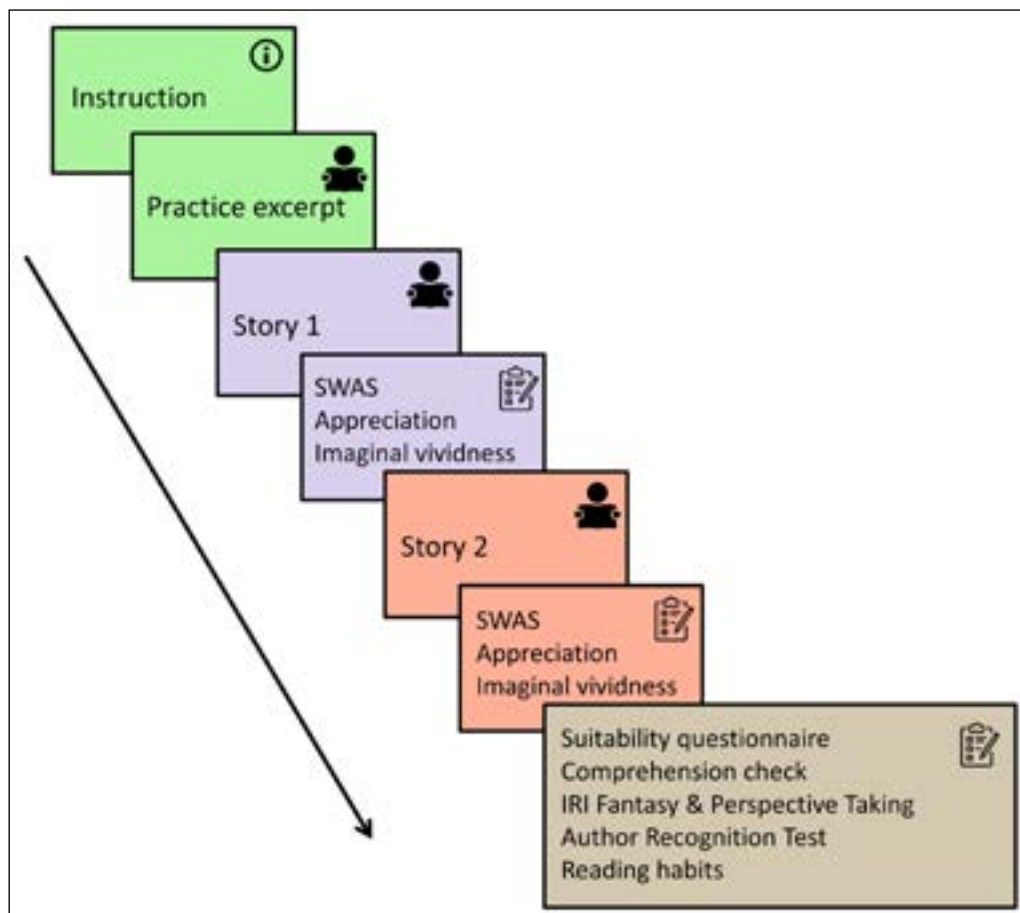
Figure 2: Schematic overview of the procedure of Experiment 2.

that were not used in experiment 1 in more detail below (see the materials section under experiment 1 for a detailed discussion of the other questionnaires).

As mentioned in the introduction of this paper we looked at appreciation of the stories. We measured this with the Appreciation Questionnaire, which is previously described in Mak & Willems (2019) and consists of a general score of story liking (*How did you like the story*; 1 = It was very bad, 7 = It was very good) and twelve adjectives (e.g., *[did you find the story] Entertaining, … Ominous*) that could be used to describe the stories (adapted from Knoop et al., 2016). These adjectives are taken from a list of adjectives that were most often used by people to describe their opinion of poetry and which are also used to describe aesthetic appeal in the domain of literature (Knoop et al., 2016). In the original scale by Mak & Willems (2019), a thirteenth adjective (*[did you find the story] Entertaining*) was used, but the Dutch word used for "entertaining" (in Dutch, "onderhoudend") appeared to be unknown to some of the participants, and was therefore removed from the questionnaire in the present study. Finally, 6 questions are asked regarding the enjoyment of the story (from Kuijpers et al., 2014; e.g., *I was constantly curious about how the story would end; I thought the story was written well*). Participants rated both the adjectives and the questions regarding enjoyment on a 7-point scale (1 = disagree, 7 = agree).

Vividness of experienced imagery was measured using a slightly adapted version of the Imaginal Vividness Scale (IVS; Fialho, personal communication), which is partly based on the Literary Response Questionnaire (Miall & Kuiken, 1995) and partly on a series of in-depth interviews with readers. This questionnaire was used because it is a more elaborate measure of imagery than the imagery subscale of the SWAS (which consists of only three items; and which is mainly focused on visual aspects of mental simulation, instead of the multisensory mental simulation we wanted to investigate, as explained in the introduction). The IVS as used in this experiment consisted of a total of 15 items divided over two subscales: Character (7 items, e.g., *While reading this story I could see the events happening in the story through the eyes of the main character; While reading the story I could almost feel the physical experiences of the characters in my own body*) and Setting (8 items, e.g., *While reading this story I often saw the described places so clearly, that it almost was as if I was there; I sometimes had auditory experiences (for example, hearing sounds) as if I was present in the world of the story*). This allowed us to capture the quality of the imagery experienced by the participants in more detail.

After reading both stories and finishing the story-related questionnaires all participants answered a suitability questionnaire asking about the suitability of the text for 14–15-year-olds and a comprehension check. The questions on the suitability questionnaire were asked in such a way, that the questionnaire would not feel "out of the blue" for participants in the control or imagery groups.[4] The suitability questionnaire was simply a follow-up on the secondary school suitability instruction, where participants had to answer three

questions about the suitability of the word use, sentence constructions, and the general suitability of the story for students in the lower grades of secondary school, of about 14 or 15 years old. The comprehension check consisted of three multiple choice questions per story (3 possible answers per question), that should have been possible to answer correctly for people who had read the stories with normal attention. Participants who answered two or more questions of the comprehension check incorrectly for one or both stories, were excluded from analysis.

Finally, participants were asked to report their reading habits and some personality characteristics. Reading habits were measured using the same questionnaire as used in the first experiment. Additionally, as an implicit measure of reading habits, participants completed the Author Recognition Test (ART; Stanovich & West, 1989; Dutch adaptation reported in Koopman, 2015), consisting of 42 names (30 real authors and 12 foils), where they had to indicate who they thought were genuine authors.

Personality characteristics were measured using the Fantasy and Perspective Taking subscales of the Interpersonal Reactivity Index (IRI; Davis, 1980; Dutch translation adapted from De Corte et al., 2007) on a 7-point scale (e.g., *Becoming extremely involved in a good book or movie is somewhat rare for me; When I'm upset at someone, I usually try to "put myself in his shoes" for a while*). The Fantasy subscale (which is conceptually related but not identical to the Transportability Scale; Dal Cin et al., 2004) measures the extent to which someone gets mentally very involved in the stories they encounter, to the point at which they imagine themselves being part of the story. The Perspective Taking subscale measures the extent to which someone is able to take another person's perspective in daily life.

Procedure

Participants were tested in small groups in lecture rooms at the university campus. One or two experimenters were always present to make sure the participants did not disturb each other and (if necessary) to answer questions. Before the start of the experiment, participants were informed about the procedure and were asked for written informed consent. At the start of the experiment, participants were given one of the three instructions (as described above; see **Figure 2** for a schematic overview of the procedure of this experiment). After having read the reading instruction and having practiced the instruction on an excerpt from an unrelated story not used in the remainder of the experiment, they started reading the first story. After reading, participants filled in the Story World Absorption Scale, the Appreciation Questionnaire, and the Imaginal Vividness scale. When they had finished, the reading instruction was repeated and participants read the second story. After finishing reading the second story participants completed the SWAS, Appreciation Questionnaire and IVS again, followed by the remaining questionnaires. The stories were read in counterbalanced order. Both the instructions and the stories were read from paper and the questionnaires were completed as paper and pencil tests. Participants were allowed to read

the stories and fill in the questionnaires at their own pace. The entire procedure took about 40 minutes.

Data-analysis

Data-analysis was done using the package *'lme4'* (Bates, Mächler, Bolker, & Walker, 2015) in R version 3.5.1 (R Core Team, 2018). We constructed a linear mixed effects regression model that predicted average scores on the SWAS based on experimental group (mental imagery instruction and secondary school suitability instruction contrasted with control). Random intercepts were allowed per participant.[5] Story (Story B contrasted with Story A), biological gender (male contrasted with female), age, self-reported reading habits, ART-score, and the Fantasy and Perspective taking subscales of the IRI were added to the model as general variables expected to explain additional variance. As a result, any effects of the experimental group would represent variance explained by the given instruction over and above variance explained by any story effects, demographic variables and other important individual difference measures. P-values were estimated using the "lmerTest" package (Kuznetsova, Brockhoff, & Christensen, 2017).

An initial effect of Experimental Group was calculated by comparing a base model (a model containing all predictors except for Experimental Group) with the full model, using an ANOVA. If this indicated a significant effect of experimental group, post-hoc pairwise comparisons (with Tukey HSD adjustment for multiple comparisons) between all three groups were made using the *'emmeans'* package in R (Lenth, 2018). All continuous predictors were centered and scaled. Variance Inflation Factors (VIFs) were calculated for all models to check for multicollinearity. All VIFs for all models were between 1 and 2, indicating that multicollinearity was not problematic in our models and all planned predictors could be entered into the models.

Similar models were constructed to predict scores on our other variables of main interest (i.e. the average scores on the IVS as a more elaborate measure of mental imagery, scores on the different components of appreciation; see **Table 2**). Additionally, we also ran analyses for the four subscales of the SWAS separately (i.e. Attention, Transportation, Emotional Engagement and Mental Imagery), and the subscales of the IVS. We analyzed the data at the level of these subscales for two reasons. Firstly, the subscales of the SWAS and IVS are arguably built up from subscales measuring diverging sub-constructs. Therefore, analyzing the subscales separately may give us additional information with regard to the processes underlying possible effects. Our second reason for analyzing the data at the level of subscales in addition to the average scores was to find out whether the null-effect for Instruction we found when looking at the variables of main interest (i.e. the average scores on the SWAS and the IVS), would also be visible within all of the subscales.

Results

Questionnaires

The overall scores on the Story World Absorption Scale as well as the scores on the four subscales of the SWAS all showed good to excellent reliability; total SWAS (18 items), $\Omega_t = .95$; Attention (5 items), $\Omega_t = .91$; Transportation (5 items), $\Omega_t = .89$; Emotional Engagement (5 items), $\Omega_t = .92$; Mental Imagery (3 items), $\Omega_t = .77$. Descriptive statistics for the questionnaire scores per subscale and per group are visualized in **Figure 3**.

The Appreciation Questionnaire was divided into two parts for the analysis. The first part, consisting of twelve adjectives that could be used to describe the stories, was analysed using a principal components analysis (PCA) with oblique rotation (direct oblimin). The Kaiser-Meyer-Olkin measure was good, KMO = .86, indicating that the sampling adequacy for this analysis was good (all KMO values for individual items >.72). Bartlett's test of sphericity showed that there was sufficient correlation between items, $\chi^2(66) = 1558.99$, $p < .001$. An initial analysis showed that three components had eigenvalues over 1 (Kaiser's criterion), but a model with three components did not fit the data well enough (fit = .93). Therefore, in the final model four components were retained (fit = .95). This

Table 2: Hierarchy of the dependent variables, including measuring instruments (and subscales) used in Experiment 2.

Construct	Measuring instrument	Variables of main interest	Subscales (if applicable)
Absorption	Story World Absorption Scale (SWAS)	1. Mean SWAS	1a. SWAS Attention
			1b. SWAS Transportation
			1c. SWAS Emotional Engagement
			1d. SWAS Mental Imagery
Vividness of mental imagery	Imaginal Vividness Scale (IVS)	2. Mean IVS	2a. IVS Character
			2b. IVS Setting
Appreciation	Appreciation Questionnaire	3a. General appreciation	
		3b. Interest	
		3c. Emotional Response	
		3d. Amusement	
		3e. Suspense	
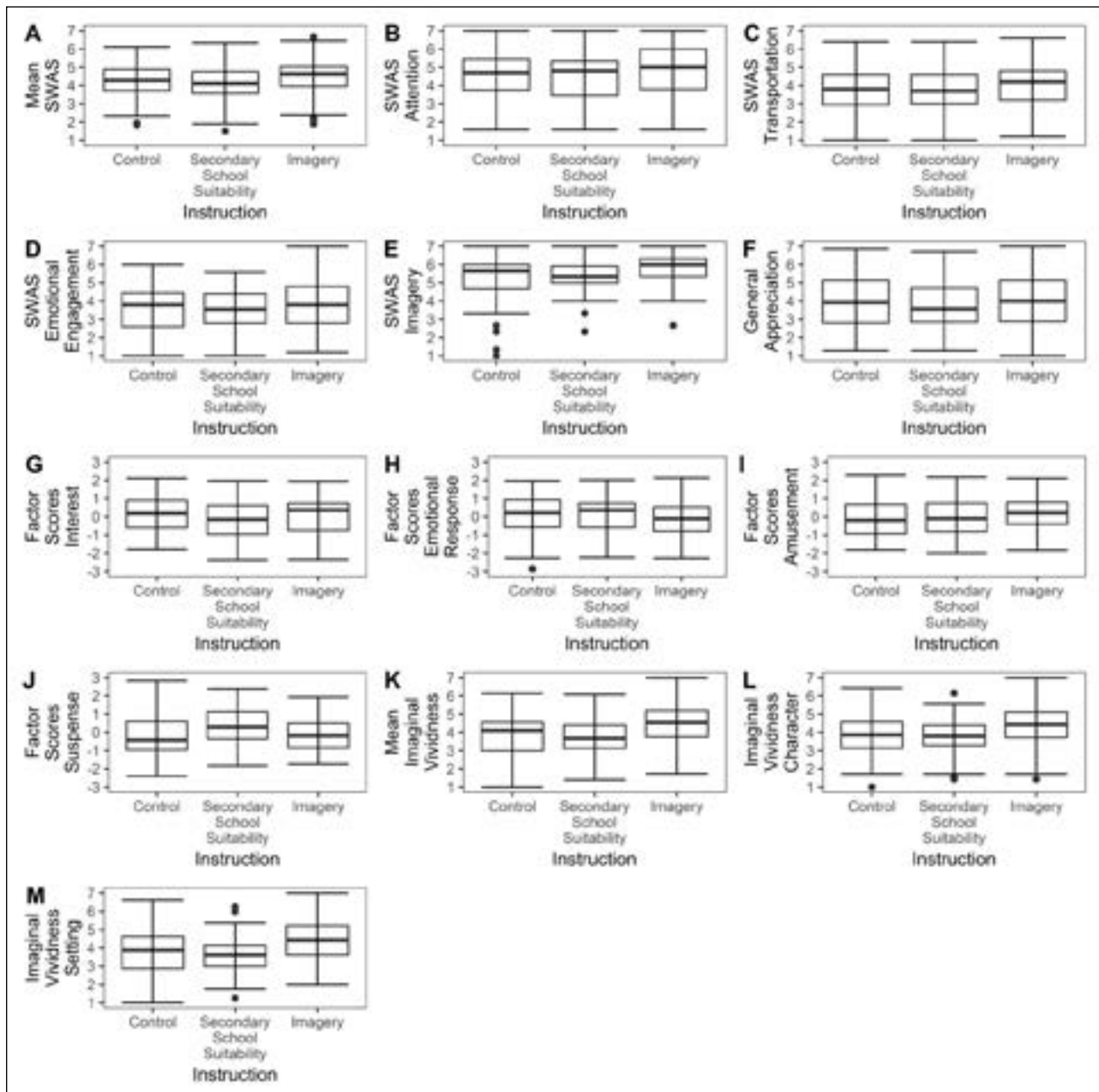
Figure 3: Boxplots depicting scores on all questionnaires and subscales, displayed per group.

model explained 71% of the variance. The first component contained items measuring the evoked interest in the story (beautiful, boring (–), interesting, captivating, deeply moving, suspenseful), the second component contained items measuring the emotional response to the story (deeply moving, tragic, sad), the third component contained items measuring the suspense elicited by the story (ominous, suspenseful), and the fourth component contained items measuring the amusement elicited by the story (funny, witty, special).[6] The structure and pattern matrices for the factor loadings after rotation can be seen in **Table 3**. Factor scores per participant and story were used in the subsequent analyses for the constructs Evoked Interest, Emotional Response, Suspense, and Amusement.

A second part of the questionnaire consisted of a general score of story liking, and 6 questions regarding the enjoyment of the story, $\Omega_t = .95$ (7 items). The answers

on these questions were collapsed into a mean score for General Appreciation.

The average scores on the Imaginal Vividness scale, as well as the scores on the two subscales, all showed sufficient to excellent reliability; total Imaginal Vividness (15 items), $\Omega_t = .93$; Character subscale (7 items), $\Omega_t = .86$; Setting subscale (8 items), $\Omega_t = .89$.

Reading experience was measured both directly using a reading habits questionnaire, and indirectly using the Author Recognition Test (ART). Because answers on different items of the reading habits questionnaire were measured on different scales, z scores were calculated for all questions on this questionnaire (higher values indicating more habitual readers). Overall reliability was sufficient, $\Omega_t = .82$. The scores on the ART were slightly positively skewed ($M = 7.46$, $SD = 4.03$, median = 7.00, IQR = 5.00–10.00) with higher values indicating more (literary) reading experience.

Table 3: Summary of the Principal Components Analysis results for the 12 adjectives on the appreciation questionnaire (N = 236). Factor loadings over .40 appear in bold.

	Structure matrix				Pattern Matrix			
	Evoked interest	**Emotional response**	**Amusement**	**Suspense**	**Evoked interest**	**Emotional response**	**Amusement**	**Suspense**
Beautiful	**0.80**	0.25	0.36	0.04	**0.81**	0.09	0.01	−0.16
Boring	**−0.82**	−0.20	**−0.42**	−0.03	**−0.81**	−0.04	−0.08	0.17
Deeply moving	**0.65**	**0.58**	0.29	0.35	**0.53**	**0.43**	−0.03	0.14
Funny	0.38	0.05	**0.87**	0.02	0.02	−0.05	**0.89**	−0.12
Interesting	**0.84**	0.28	**0.55**	0.31	**0.71**	0.06	0.20	0.10
Ominous	0.09	**0.41**	0.19	**0.86**	−0.20	0.23	0.11	**0.83**
Sad	0.26	**0.85**	0.11	0.19	0.08	**0.85**	−0.05	−0.04
Suspenseful	**0.55**	0.09	0.20	**0.70**	**0.50**	−0.18	−0.11	**0.66**
Tragic	0.23	**0.87**	0.14	0.29	0.01	**0.85**	0.01	0.07
Witty	**0.49**	0.19	**0.82**	0.34	0.11	0.01	**0.73**	0.20
Captivating	**0.87**	0.28	**0.54**	0.38	**0.75**	0.03	0.17	0.18
Special	**0.58**	0.28	**0.60**	0.36	0.33	0.10	**0.40**	0.20

Reliability of both subscales of the Interpersonal Reactivity Index was sufficient; Fantasy subscale (7 items), $\Omega_t = .84$, and Perspective Taking subscale (7 items), $\Omega_t = .84$.

Main Analysis

The model predicting average scores on the SWAS based on story, biological gender, age, ART-score, reading habits, IRI Fantasy, IRI Perspective Taking and experimental group showed no differences in SWAS scores between the three experimental groups (see **Figure 3A**; see **Table 4.1** for all results of this model). Interestingly, participants with higher scores on the Fantasy subscale of the IRI reported more story world absorption (see **Figure 4A**). To find out which aspects of story world absorption this relationship between IRI fantasy and scores on the SWAS stems from, similar models were constructed to predict the scores on the four subscales of the SWAS (see **Figure 4** for a visual representation of the relationship between scores on the Fantasy subscale of the IRI and the tested reading experiences).

The positive relationship between scores on the Fantasy subscale of the IRI and scores on the Story World Absorption Scale was visible on all subscales of the SWAS (see **Figure 4B–E**; see **Table 4.1a–d**). On the Attention and Mental Imagery subscales, story effects became visible; after reading Story B participants reported higher attention to the story world and higher Mental Imagery. In contrast, participants reported lower emotional engagement with Story B (see **Table 4.1a–d**). No other significant associations between the predictors and any of the subscales of the SWAS were found.

To investigate the effect of mental imagery instructions on reported mental imagery more thoroughly, we also tested differences between groups in mental imagery as reported on the IVS (see **Figure 3K**; see **Table 4.2**). The results on this questionnaire also indicate whether participants complied with our mental imagery instructions. A comparison between our base model and the full model suggested a significant effect of group on scores on the IVS ($F(2, 118) = 6.49$, $p = .002$), but post-hoc comparison of the two experimental groups and the control group showed no notable differences (Mental Imagery Instruction vs. Control: $B = −0.35$, $SE = 0.19$, $df = 128$, $t = −1.90$, $p = .14$; Secondary School Suitability instruction vs. Control: $B = 0.29$, $SE = 0.18$, $df = 128$, $t = 1.56$, $p = .27$). However, the group receiving the Mental Imagery instruction did report significantly more vivid imagery than the group receiving the Secondary School Suitability instruction ($B = 0.64$, $SE = 0.19$, $df = 128$, $t = 3.46$, $p = .002$). IRI Fantasy was positively related to the vividness of mental imagery (see **Figure 4K**). After reading Story B, participants reported more vivid imagery. To find out if the effect of instruction was perhaps only visible on one of the two subscales and to find out which aspects of imaginal vividness the relationship between IRI fantasy and imaginal vividness stems from, these analyses were repeated for the individual subscales of the IVS.

The Character subscale of the IVS showed similar results for the relationship between instructions and imaginal vividness as were found on the overall scale: An initial comparison between our base model and the full model suggested a significant effect of type of instruction ($F(2, 118) = 3.61$, $p = .03$; see **Figure 3L**), but post-hoc comparison of the two experimental groups and the control group did not reveal any statistically significant differences (Mental Imagery instruction vs. Control: $B = −0.30$, $SE = 0.20$, $df = 128$, $t = −1.47$, $p = .31$; Secondary School Suitability instruction vs. Control: $B = 0.22$, $SE = 0.20$, $df = 128$, $t = 1.10$, $p = .51$). The group receiving the Mental Imagery instruction did score higher on the Character subscale of the IVS than the group receiving the Secondary School Suitability instruction ($B = 0.51$, $SE = 0.20$, $df = 128$, $t = 2.58$, $p = .03$).

Table 4: Coefficients of the models predicting reading experiences (absorption, vividness of mental imagery, story appreciation) based on type of instruction (mental imagery instruction and secondary school suitability instruction contrasted with the control instruction), taking into account Story (Story B contrasted with Story A), biological gender (male contrasted with female), age, self-reported reading habits, ART-score, and the Fantasy and Perspective taking subscales of the IRI. Significant predictors are marked (* $p < .05$, ** $p < .01$, *** $p < .001$).

		B	SE	df	t-value	p-value
1. SWAS	(Intercept)	4.22	0.13	156.82	32.37	<.001***
	Imagery instruction	0.17	0.16	118.00	1.01	.32
	School instruction	−0.12	0.16	118.00	−0.71	.48
	IRI Fantasy	0.39	0.08	118.00	5.12	<.001***
	IRI Perspective taking	−0.04	0.07	118.00	−0.60	.55
	Reading habits	0.09	0.13	118.00	0.70	.49
	ART-score	−0.09	0.08	118.00	−1.09	.28
	Age	0.06	0.08	118.00	0.77	.44
	Biological gender (male)	0.30	0.19	118.00	1.60	.11
	Story (B)	0.03	0.10	118.00	0.28	.78
1a. SWAS Attention	(Intercept)	4.42	0.17	149.38	25.89	<.001***
	Imagery instruction	0.10	0.22	118.00	0.48	.64
	School instruction	−0.22	0.22	118.00	−1.00	.32
	IRI Fantasy	0.32	0.10	118.00	3.13	.002**
	IRI Perspective taking	−0.02	0.09	118.00	−0.22	0.83
	Reading habits	0.31	0.17	118.00	1.81	.07
	ART-score	−0.13	0.11	118.00	−1.14	.26
	Age	0.03	0.10	118.00	0.32	.75
	Biological gender (male)	0.28	0.25	118.00	1.12	.27
	Story (B)	0.47	0.12	118.00	3.95	<.001***
1b. SWAS Transportation	(Intercept)	3.65	0.16	147.16	22.12	<.001***
	Imagery instruction	0.16	0.21	118.00	0.74	.46
	School instruction	−0.07	0.21	118.00	−0.34	.74
	IRI Fantasy	0.49	0.10	118.00	4.98	<.001***
	IRI Perspective taking	−0.09	0.09	118.00	−1.00	.32
	Reading habits	0.13	0.17	118.00	0.79	.43
	ART-score	−0.14	0.11	118.00	−1.33	.19
	Age	0.11	0.10	118.00	1.10	.28
	Biological gender (male)	0.45	0.25	118.00	1.85	.07
	Story (B)	0.15	0.11	118.00	1.39	0.17
1c. SWAS Emotional Engagement	(Intercept)	4.08	0.16	165.90	25.11	<.001***
	Imagery instruction	0.17	0.20	118.00	0.86	.39
	School instruction	−0.07	0.20	118.00	−0.38	.71
	IRI Fantasy	0.38	0.09	118.00	4.02	<.001***
	IRI Perspective taking	−0.03	0.08	118.00	−0.32	.75
	Reading habits	−0.06	0.16	118.00	−0.39	.70
	ART-score	−0.08	0.10	118.00	−0.78	.44
	Age	0.01	0.09	118.00	0.07	.94

(Contd.)

		B	SE	df	t-value	p-value
1d. SWAS Mental Imagery	Biological gender (male)	0.22	0.23	118.00	0.95	.35
	Story (B)	−0.89	0.14	118.00	−6.55	<.001***
	(Intercept)	5.08	0.12	158.85	41.03	<.001***
	Imagery instruction	0.27	0.16	118.00	1.75	0.08
	School instruction	−0.09	0.15	118.00	−0.60	0.55
	IRI Fantasy	0.38	0.07	118.00	5.20	<.001***
	IRI Perspective taking	−0.02	0.06	118.00	−0.30	.76
	Reading habits	−0.09	0.12	118.00	−0.77	.44
	ART-score	0.04	0.08	118.00	0.50	.62
	Age	0.11	0.07	118.00	1.50	.14
2. IVS	Biological gender (male)	0.23	0.18	118.00	1.29	.20
	Story (B)	0.62	0.10	118.00	6.39	<.001***
	(Intercept)	3.82	0.14	143.53	27.73	<.001***
	Imagery instruction	0.35	0.18	118.00	1.97	.051
	School instruction	−0.29	0.18	118.00	−1.63	.11
	IRI Fantasy	0.47	0.08	118.00	5.69	<.001***
	IRI Perspective taking	−0.08	0.07	118.00	−1.06	.29
	Reading habits	0.12	0.14	118.00	0.86	.39
	ART-score	−0.11	0.09	118.00	−1.15	.25
	Age	0.07	0.08	118.00	0.91	.36
2a. IVS Character	Biological gender (male)	−0.22	0.21	118.00	−1.09	.28
	Story (B)	0.33	0.09	118.00	3.77	<.001***
	(Intercept)	3.94	0.15	145.97	26.31	<.001***
	Imagery instruction	0.30	0.19	118.00	1.53	.13
	School instruction	−0.22	0.19	118.00	−1.15	0.25
	IRI Fantasy	0.45	0.09	118.00	5.02	<.001***
	IRI Perspective taking	−0.10	0.08	118.00	−1.23	.22
	Reading habits	0.07	0.15	118.00	0.46	.65
	ART-score	−0.13	0.10	118.00	−1.37	.17
	Age	0.03	0.09	118.00	0.33	.75
2b. IVS Setting	Biological gender (male)	−0.29	0.22	118.00	−1.30	.20
	Story (B)	0.21	0.10	118.00	2.12	.04*
	(Intercept)	3.73	0.14	145.24	26.15	<.001***
	Imagery instruction	0.40	0.18	118.00	2.15	.03*
	School instruction	−0.35	0.18	118.00	−1.93	.06
	IRI Fantasy	0.49	0.09	118.00	5.75	<.001***
	IRI Perspective taking	−0.06	0.08	118.00	−0.78	.44
	Reading habits	0.17	0.14	118.00	1.15	0.25
	ART-score	−0.08	0.09	118.00	−0.86	.39
	Age	0.11	0.08	118.00	1.35	.18
	Biological gender (male)	−0.17	0.21	118.00	−0.79	.43
	Story (B)	0.43	0.09	118.00	4.62	<.001***

(Contd.)

		B	SE	df	t-value	p-value
3a. General Appreciation	(Intercept)	3.55	0.19	162.42	18.74	<.001***
	Imagery instruction	−0.14	0.24	118.00	−0.60	.55
	School instruction	−0.22	0.23	118.00	−0.94	.35
	IRI Fantasy	0.36	0.11	118.00	3.23	.002**
	IRI Perspective taking	−0.03	0.10	118.00	−0.32	.75
	Reading habits	0.07	0.19	118.00	0.37	.71
	ART-score	−0.05	0.12	118.00	−0.44	.66
	Age	0.13	0.11	118.00	1.18	.24
	Biological gender (male)	0.36	0.27	118.00	1.30	.19
	Story (B)	0.84	0.15	118.00	5.51	<.001***
3b. Factor Scores Interest	(Intercept)	−0.11	0.13	166.17	−0.84	.40
	Imagery instruction	−0.19	0.16	117.67	−1.17	.24
	School instruction	−0.32	0.16	118.67	−1.97	.052
	IRI Fantasy	0.18	0.08	117.78	2.39	.02*
	IRI Perspective taking	−0.02	0.07	118.39	−0.30	.77
	Reading habits	0.15	0.13	117.72	1.16	.25
	ART-score	−0.06	0.08	117.67	−0.74	.46
	Age	0.09	0.07	117.79	1.14	.26
	Biological gender (male)	0.29	0.19	120.55	1.51	.13
	Story (B)	0.46	0.11	117.98	4.12	<.001***
3c. Factor Scores Emotional Response	(Intercept)	−0.10	0.14	115.01	−0.69	.49
	Imagery instruction	−0.28	0.18	116.44	−1.58	.12
	School instruction	−0.03	0.18	117.38	−0.17	.86
	IRI Fantasy	0.15	0.08	116.55	1.85	.07
	IRI Perspective taking	0.02	0.07	117.12	0.28	.78
	Reading habits	−0.02	0.14	116.49	−0.15	.88
	ART-score	0.04	0.09	116.44	0.40	.69
	Age	−0.004	0.08	116.56	−0.05	.96
	Biological gender (male)	0.28	0.21	119.15	1.34	.18
	Story (B)	0.30	0.11	116.61	2.80	.006**
3d. Factor Scores Amusement	(Intercept)	−0.38	0.13	149.28	−2.89	.004**
	Imagery instruction	0.17	0.17	117.76	1.01	.31
	School instruction	0.02	0.17	118.63	0.10	.92
	IRI Fantasy	0.29	0.08	117.86	3.60	<.001***
	IRI Perspective taking	−0.05	0.07	118.39	−0.78	.43
	Reading habits	−0.02	0.13	117.80	−0.13	.90
	ART-score	0.04	0.09	117.76	0.51	.61
	Age	−0.16	0.08	117.86	−2.08	.04*
	Biological gender (male)	0.31	0.20	120.27	1.56	.12
	Story (B)	0.55	0.09	117.80	5.90	<.001***
3e. Factor Scores Suspense	(Intercept)	−0.36	0.14	142.78	−2.59	.01*
	Imagery instruction	0.004	0.18	117.35	0.02	.98
	School instruction	0.44	0.18	118.14	2.45	.02

(Contd.)

	B	*SE*	*df*	*t*-value	*p*-value
IRI Fantasy	0.11	0.09	117.44	1.27	.21
IRI Perspective taking	−0.11	0.07	117.92	−1.52	.13
Reading habits	−0.14	0.14	117.39	−0.95	.34
ART-score	−0.01	0.09	117.35	−0.09	.93
Age	−0.12	0.08	117.45	−1.39	.17
Biological gender (male)	0.01	0.21	119.63	0.03	.97
Story (B)	0.44	0.09	117.26	4.97	<.001***
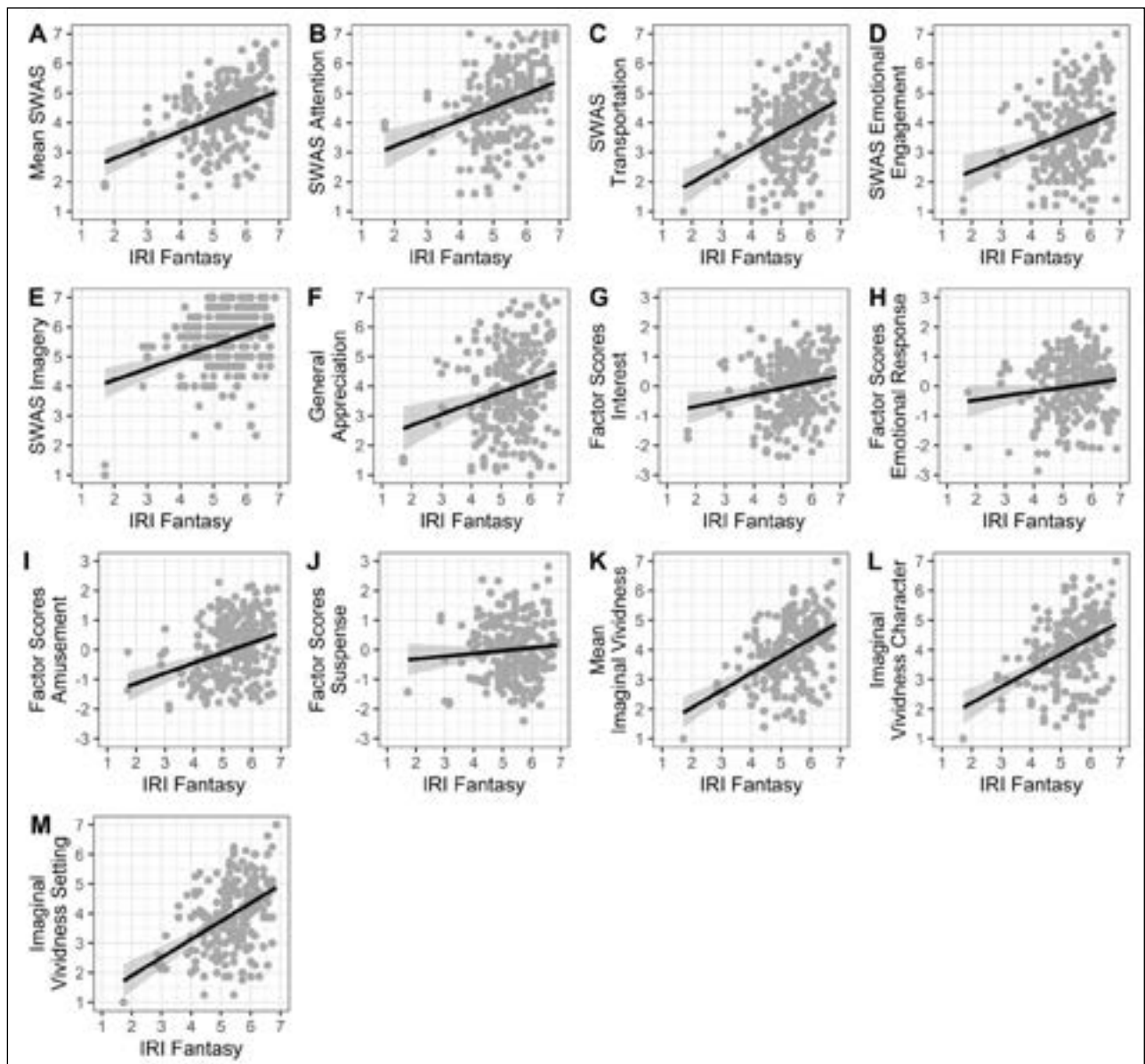
Figure 4: Relationship between scores on the Fantasy subscale of the IRI and the tested reading experiences.

The same pattern was revealed for the Setting subscale of the IVS: an initial comparison between our base model and the full model suggested a significant effect of type of instruction ($F(2, 118) = 8.35$, $p < .001$; see **Figure 3M**), but post-hoc comparisons of the two experimental groups and the control group did not reveal any statistically significant differences (Mental Imagery instruction vs. Control: $B = -0.40$, $SE = 0.19$, $df = 128$, $t = -2.07$, $p = .10$;

Secondary School Suitability instruction vs. Control: $B = 0.35$, $SE = 0.19$, $df = 128$, $t = 1.86$, $p = .16$). Again, the group receiving the Mental Imagery instruction scored higher on the Setting subscale of the IVS than the group receiving the Secondary School Suitability instruction ($B = 0.75$, $SE = 0.19$, $df = 128$, $t = 3.93$, $p < .001$).

On both subscales of the Imaginal Vividness Scale, we found a positive relationship between scores on the

Fantasy subscale of the IRI and imaginal vividness (see **Figure 4L&M**; See **Table 4.2a–b**). Similarly, differences between stories were found for both subscales of the IVS. After reading Story B, participants reported more vivid imagery of the characters in the story and of the settings described in the story.

The results on the Imaginal Vividness Scale suggest that our reading instructions indeed influenced the experienced vividness of mental imagery, with respect to both the characters in the story and the environment described in the stories. The imagery instruction was associated with more vivid mental imagery than the secondary school suitability instruction. This suggests that participants are able to follow these instructions while reading, and that they indeed target mental imagery, as intended. Apart from instructed mental imagery, we also found a significant role for the personality trait Fantasy in the experienced vividness of mental imagery.

To test whether mental imagery instructions would also have an impact on the appreciation of stories, we tested differences between groups in general appreciation and story appreciation as reported on the four components of the Appreciation Questionnaire (see **Table 4.3a–e**). Initial comparisons between the base models and the full models showed that the reading instructions influenced only the experienced Suspense ($F(2, 117.90) = 3.95$, $p = .02$; see **Figure 3J**), although post-hoc comparisons of the three groups failed to reach statistical significance (Mental Imagery instruction vs. Control: $B = -0.004$, $SE = 0.19$, $df = 127$, $t = -0.02$, $p = .9998$; Secondary School Suitability instruction vs. Control: $B = -0.44$, $SE = 0.19$, $df = 128$, $t = -2.35$, $p = .052$; Mental Imagery Instruction vs. Secondary School Suitability instruction: $B = -0.44$, $SE = 0.19$, $df = 128$, $t = -2.32$, $p = .057$; Note however, that – although not statistically significant – this suggests that the participants receiving the secondary school suitability instructions experienced the stories they read as being somewhat more ominous and suspenseful, opposite to our expectations). The reading instructions did not have an effect on the four other aspects of appreciation.

Comparable to our findings for the SWAS, participants scoring higher on the Fantasy subscale of the IRI appreciated the stories they read more (General Appreciation; see **Figure 4F**; see **Table 4.3a**). Similarly, a positive association was found between IRI Fantasy and factor scores for Evoked Interest (see **Figure 4G**; see **Table 4.3b**) and the factor scores for Amusement (see **Figure 4I**; see **Table 4.3d**). Additionally, there was a negative association between age and factor scores for Amusement: older participants reported finding the stories less funny, witty or special (see **Table 4.3d**). Differences between stories were found for General Appreciation, Evoked Interest, Emotional Response, Amusement, and Suspense (see **Table 4.3a–e**). Story B was generally appreciated more, evoked more interest, elicited a stronger emotional response, was considered more Amusing, and more Suspenseful.

To test whether there was a moderation between instruction condition and IRI Fantasy, we performed exploratory analyses where we included an interaction term between IRI Fantasy and instruction in our models. There was a significant interaction between IRI Fantasy and the Secondary School Suitability instruction for three of the thirteen tested dependent variables (SWAS Mental Imagery: $B = -0.36$, $SE = 0.14$, $df = 118.00$, $t = -2.91$, $p = .004$; Emotional Response: $B = -0.47$, $SE = 0.16$, $df = 116.5$, $t = -2.91$, $p = .004$; Suspense: $B = 0.38$, $SE = 0.17$, $df = 117.40$, $t = 2.28$, $p = .02$). From the visualization of these interactions in **Figure 5** can be seen that for Mental Imagery and Emotional Response, the relationship between IRI Fantasy and the dependent variable is attenuated when participants have to read with a reading instruction in mind (and mostly so if this is the Secondary School Suitability instruction). Oppositely, the relationship between IRI Fantasy and Emotional Response seems only present in participants who received the Secondary School Suitability instruction, but not in the other groups. Although these results are interesting in itself, we have to be careful with interpreting them, as this moderation only appears for a few of our dependent variables and does not follow a highly consistent pattern.

Discussion

The aim of this second experiment was to replicate the findings of the first experiment in a more controlled setting, with additional stories, and while considering an extra set of control variables (most importantly aiming at personality characteristics that might influence reading experiences). As can be seen from the results of this experiment, reading instructions only played a very minor role in defining reading experiences.[7] Although we could see that, in particular, our Mental Imagery instruction did influence the reading experiences directly involving mental imagery, suggesting that our instruction was indeed successful in influencing mental imagery, the effect of this instruction did not translate to other reading experiences.[8]

The most notable statistically significant finding in this study was that the Fantasy subscale of the Interpersonal Reactivity Index appeared to be positively associated with all aspects of participants' reading experiences. Even though it may be possible to influence certain reading experiences through reading instructions, personality characteristics appear to be much more important in determining people's reading experiences. As was described above, the Fantasy subscale measures the extent to which someone has the tendency to get mentally involved in the stories they encounter by imagining themselves being part of the story or by trying to empathize with characters in the story. Together, the questions on this subscale of the IRI give an impression of the amount of "fantasy" with which participants experience fiction on a day to day basis. Because of the theoretical relationship between this personality characteristic and reading experiences, it is interesting to find that this personality characteristic is indeed positively associated with reading experiences across the board (note that this personality characteristic is not just associated with absorption or mental imagery, but also with several aspects of how participants appreciated the stories). A question of causality remains, however: future studies will have to determine whether those who engage more in imagery during reading, as a consequence like the stories they read more, or whether people who
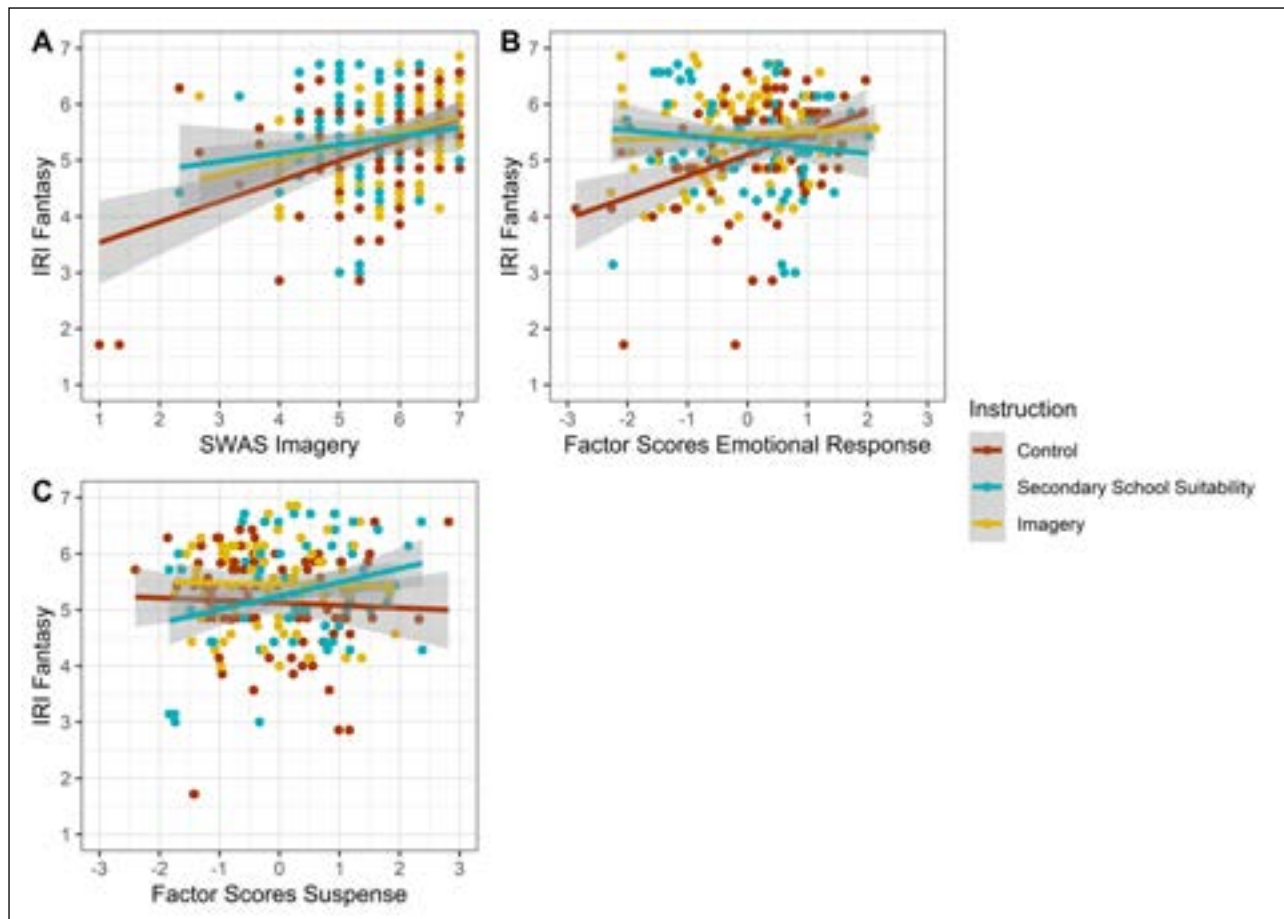
Figure 5: Interaction between IRI Fantasy and instruction group, for SWAS Mental Imagery scores (5A), and factor scores for Emotional Response (5B) and Suspense (5C).

enjoy stories gradually become more imaginative as a result of reading (comparable to the question of causality in the study of the relationship between reading and theory of mind: do more empathic people read more, or does reading result in more empathy? See Panero et al., 2016; Samur, Tops, & Koole, 2017).

Finally, it is interesting to find that the influences of biological gender and reading habits on reading experiences as found in the first experiment reported in this paper were not significantly associated with reading experiences in the second experiment. Just like the influence of reading instructions, the effects of biological gender and reading habits do not seem to be as important as personality characteristics in determining reading experiences (although note that the variation in biological gender in this second experiment was far from balanced: in the second experiment only ~20% of participants were male, compared to ~50% in experiment 1. Therefore, no strong conclusions about the effects of biological gender can be drawn from the results of experiment 2). As opposed to the results in our previous study, age was negatively associated with amusement: older participants found the stories less funny, witty or special. However, although interesting in itself, it should be noted that the majority of participants in this study were university students of about 21 or 22 years of age, and this effect might be due to a couple of outliers (only 7 participants were older than 30 years of age, with 3 of them being 55 years or older). However, a previous study with a larger variation in age

between participants (and a large number of participants being between 50 and 75 years old) also showed that older participants rated the stories they read as less literary and less beautiful than younger participants did (Hartung et al., 2017). Our results showed a few differences between the two stories with respect to the reading experiences they elicited. As we did not have any hypotheses regarding how the stories would differ, we will not interpret the results regarding the differences between the stories.

General Discussion

In this study we investigated the relationship between mental imagery and reading experiences. In particular, we were interested in the act of mental imagery during reading and whether differences between people in the extent to which they engaged in mental imagery was related to their reading experiences. To make sure that participants differed in the extent to which they engaged in mental imagery, they received reading instructions in which they were instructed to envision the stories as much as possible, to read as if they were reading for leisure, or to focus on surface characteristics of the stories (word use and sentence construction; only in experiment 2). Apart from mental imagery, we were interested in the role of stable or trait-like individual differences (such as reading habits, biological gender, age, education, and personality characteristics) in determining reading experiences.

Although experiment 1 suggested that mental imagery instructions, as well as level of education, biological

gender, and reading habits, played a significant role in determining reading experiences, experiment 2 showed that after controlling for personality characteristics (in particular "fantasy") and adding an extra control condition, this association between mental imagery instructions, biological gender, reading habits, and reading experiences disappeared for a large part. This suggests that, besides all other aspects involved in reading experiences, these experiences are most strongly influenced by personality characteristics, such as readers' proneness to "fantasy". Fantasy has been suggested to be one of the aspects underlying the "Openness" personality characteristic (Fayn et al., 2015), a characteristic that has also been found to be associated to reading experiences and reading habits in other studies (Kraaykamp & Eijck, 2005; Kuijpers et al., 2018; Schutte & Malouff, 2004).

Mental imagery was mainly found to be related to mental imagery-related reading experiences, and not as strongly to other reading experiences. However, in a previous study mental simulation was found to be related to aspects of absorption and appreciation (Mak & Willems, 2019). The reason that this relationship was not found to be very strong in the current study, could be that there is a difference between (explicit) mental imagery and (implicit) mental simulation (see Jacobs & Willems, 2018). Perhaps the explicit mental imagery the participants were instructed to perform in this study was too different from the implicit mental simulation elicited by stories during naturalistic reading, and was therefore relatively unrelated to reading experiences. This could also explain why submitting participants to a more implicit mental imagery training before reading, did prove effective in increasing experienced transportation (Johnson et al., 2013).

Moreover, the interactions between the effects of fantasy and reading instructions on some of the tested reading experiences in experiment 2 even suggest that pre-reading instruction might in fact negatively influence naturalistic processes during reading. For both mental imagery as reported on the SWAS and for the emotional response to the story, it was found that the positive relationship between fantasy and these reading experiences in readers in the control group, was attenuated in readers who received pre-reading instructions (regardless of the content of the instructions). Therefore, it could be possible that having to remember and execute instructions during reading interferes with reading experiences as they would normally occur. However, the interactions that were found were not present for all aspects of reading experiences, and these analyses were highly exploratory, so further research should indicate whether this is indeed the case. However, when studying subjective reading experiences, it seems wise to only study naturalistic reading instead of trying to influence reading experiences using pre-reading instructions.

Another explanation for the weak association between mental imagery and reading experiences in this study could be that readers differ greatly in the type of mental imagery they prefer during reading. Kuzmičová (2014) suggests four possible types of mental imagery during literary reading: Enactment-imagery (where readers imagine themselves executing the actions described in the story), description-imagery (where readers visually imagine the objects and scenes described in the story), speech-imagery (where readers imagine hearing the narrator tell the story) and rehearsal-imagery (where readers imagine reading the story out lout). Kuzmičová suggests readers differ in the type of imagery they perform during reading (and this can also differ from one story to the next within a given reader). Perhaps the instructions given in this experiment did not match the preferred type of imagery of some (or all) of the readers, resulting in weak effects of our mental imagery instruction on reading experiences.

A different possibility would be that mental imagery just doesn't play a role in people's ability to become involved in a story. However, previous research has shown relationships between imagery and absorption, transportation, and appreciation or enjoyment of narratives, which does not fit with the proposal that imagery is unimportant in story involvement (Green, 2004; Green & Brock, 2002; Kuijpers et al., 2014; Mol & Jolles, 2014; Weibel et al., 2011). Another possibility would be that people are unable to perform mental imagery "on command". However, the fact that reading instructions were successful at inducing or reducing mental imagery in our participants contradicts this claim.

Overall, it seems that the use of mental imagery-inducing reading instructions does have a small influence on people's reading experiences. However, this effect pales into insignificance compared to the effect of personality characteristics. Perhaps a single reading instruction is insufficient for altering reading experiences: to really enhance reading experiences, readers will have to be trained intensively to read in a different way. For instance, Janssen, Braaksma and Couzijn (2009) found that students' appreciation for stories increased after an intervention where they were allowed to come up with their own questions about the stories (as opposed to answering a teacher's questions). Perhaps a comparable intervention, but instead aimed at mental imagery, might have a stronger influence on reading experiences than a single instruction, and may perhaps prove more powerful in overcoming personality characteristics (but see De Koning & van der Schoot, 2013).

Data Accessibility Statement

The participant data and analysis scripts can be found on this paper's project page on the Open Science Framework (osf.io/98ntg/). The stories used as materials in this experiment are from professional authors and are copyrighted.

Notes

[1] Mental imagery during reading is also sometimes referred to as mental simulation. Theoretically, mental simulation is a somewhat more subconscious process than mental imagery. For the sake of clarity, we will call the process of imagining events or perceptible elements of the story world described in a story (more or less consciously) mental imagery throughout this article.

[2] Note that when looking at the literature regarding mental imagery during reading it becomes clear that people seem not just to mentally image descriptions of *visual* elements of story worlds, but a much more extensive range of perceptual descriptions (i.e. auditory, olfactory, proprioceptive) and motor descriptions (Kuzmičová, 2012; Kuzmičová, 2014; Mak & Willems, 2019; Nijhof & Willems, 2015). Although the exact nature of mental imagery during reading is still debated, in this paper we do not want to focus on the *content* of mental imagery during reading, but instead on the *act* of mental imagery itself.

[3] Note however, that Kuiken and Douglas (Kuiken & Douglas, 2017) proposed that absorption is even more multidimensional, suggesting multiple types of absorption, with multiple different outcomes.

[4] Some participants still found these questions "out of the blue". However, the majority of the participants did not report being surprised by these questions. Furthermore, this questionnaire was only presented at the end of the experiment, after all the questionnaires regarding reading experiences had already been filled in, and will therefore not have affected answers on the questionnaires that were of interest for our analyses.

[5] We did not have enough observations in our dataset to support any random slopes (let alone a maximal model with random slopes for all predictor variables). Therefore, we decided that it was most appropriate to use a random intercept only model.

[6] Note that the items "deeply moving" and "suspenseful" load strongly (above .40) on more than one component. This is due to the nature of PCA as an unsupervised dimension reduction method. Items loading strongly on a component are considered "typical" items for the component, and can be used for the interpretation of the components. However, every item will load on every component, and the loadings of all items (both items strongly associated with the components and items weakly associated with the components) are taken into account when calculating the component or factor scores of all components. The strong association of the items "deeply moving" and "suspenseful" with more than one component therefore does not have important consequences for the calculation of the component scores, but only for the theoretical interpretation of the components.

[7] A possible concern with our analyses would be that the lack of effect of reading instructions is due to the other predictors (The Fantasy and Perspective Taking subscales of the IRI, scores on the Author Recognition Test) absorbing so much variance that effects of reading instructions would not become visible. To check this, we ran reduced models that are comparable to the models used in Experiment 1 (without extra predictors, but including a random intercept for Participant and a predictor for Story, to control for participant and story effects). Although there were some minor changes in the effect of reading instruction on some of the dependent variables (i.e. SWAS Mental Imagery, Mean Imaginal Vividness, the Setting subscale of the Imaginal Vividness scale, and the Suspense component of the Appreciation Questionnaire), this concerned more pronounced effects rather than effects we failed to find with the models that did include the additional measures. As the latter models were significantly better than the reduced models (based on AIC, BIC and LogLikelihood), we chose to report only the results for the complete models.

[8] Note that it is possible that the findings with regard to the vividness of mental imagery are due to experimental demand, as the mental imagery instruction specifically asked participants to increase their mental imagery. However, we think this not the most likely explanation of our results. If the imagery findings would be entirely due to the experimental demand, we would have expected that both the secondary school instruction group *and* the control group would differ from the mental imagery instruction group with respect to reported vividness of mental imagery (since neither the secondary school suitability instruction nor the control instruction mentioned mental imagery). In contrast, the situation is that the secondary school suitability instruction lowered reported mental imagery somewhat with respect to the control group, and the mental imagery instruction increased reported mental imagery somewhat, resulting in a significant difference between the secondary school suitability group and the mental imagery instruction group (with the control group being somewhere in between). Although it is possible that the decrease in mental imagery in the secondary school suitability group is coincidental, and the increase in mental imagery in the imagery instruction group is due to experimental demand, it seems more likely that both instructions had a (small) influence on mental imagery – but in opposite directions. Importantly, our general conclusions remain the same, regardless of which explanation of the results is true.

Competing Interests
No potential competing interests are reported by the authors.

RW is an editor at Collabra: Psychology and was not involved in the peer review of this article.

Author Contributions
- Contributed to conception and design: MM, CdV, RW
- Contributed to acquisition of data: MM, CdV
- Contributed to analysis and interpretation of data: MM, CdV, RW

· Drafted and revised the article: MM, RW
· Approved the submitted version for publication: MM, CdV, RW

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*(1), 278–289. DOI: https://doi.org/10.3758/BRM.40.1.278

Appel, M., & Richter, T. (2010). Transportation and need for affect in narrative persuasion: A mediated moderation model. *Media Psychology*, *13*(2), 101–135. DOI: https://doi.org/10.1080/15213261003799847

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Biesheuvel, J. M. A. (1972). *In de bovenkooi*. Amsterdam, The Netherlands: Meulenhoff.

Bilandzic, H., & Busselle, R. W. (2011). Enjoyment of films as a function of narrative experience, perceived realism and transportability. *Communications*, *36*, 29–50. DOI: https://doi.org/10.1515/comm.2011.002

Brysbaert, M., Sui, L., Dirix, N., & Hintz, F. (2020). Dutch Author Recognition Test. *Journal of Cognition*, *3*(1), 1–14. DOI: https://doi.org/10.5334/joc.95

Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, *18*, 255–280. DOI: https://doi.org/10.1111/j.1468-2885.2008.00322.x

Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychology*, *12*(4), 321–347. DOI: https://doi.org/10.1080/15213260903287259

Chiu, M. M., & McBride-Chang, C. (2006). Gender, context, and reading: A comparison of students in 43 countries. *Scientific Studies of Reading*, *10*(4), 331–362. DOI: https://doi.org/10.1207/s1532799xssr1004_1

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row. DOI: https://doi.org/10.1017/CBO9781107415324.004

Dal Cin, S., Zanna, M. P., & Fong, G. T. (2004). Narrative persuasion and overcoming resistance. In E. S. Knowles & J. A. Linn (Eds.), *Resistance and persuasion* (pp. 175–191). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers. Retrieved from https://www.taylorfrancis.com/books/9781410609816. DOI: https://doi.org/10.1037/e633872013-232

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*, 85.

De Corte, K., Buysse, A., Verhofstadt, L. L., Roeyers, H., Ponnet, K., & Davis, M. H. (2007). Measuring empathic tendencies: Reliability and validity of the Dutch version of the Interpersonal Reactivity Index. *Psychologica Belgica*, *47*(4), 235–260. DOI: https://doi.org/10.5334/pb-47-4-235

De Koning, B. B., & van der Schoot, M. (2013). Becoming part of the story! Refueling the interest in visualization strategies for reading comprehension. *Educational Psychology Review*, *25*, 261–287. DOI: https://doi.org/10.1007/s10648-013-9222-6

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*, 399–412. DOI: https://doi.org/10.1111/bjop.12046

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. DOI: https://doi.org/10.3758/BRM.41.4.1149

Fayn, K., Tiliopoulos, N., & MacCann, C. (2015). Interest in truth versus beauty: Intellect and Openness reflect different pathways towards interest. *Personality and Individual Differences*, *81*, 47–52. DOI: https://doi.org/10.1016/j.paid.2014.12.031

Fialho, O. (n.d.). Transformative Reading Scale (Manuscript in preparation).

Fong, K., Mullin, J. B., & Mar, R. A. (2013). What you read matters: The role of fiction genre in predicting interpersonal sensitivity. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(4), 370–376. DOI: https://doi.org/10.1037/a0034084

Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*. Boston: Yale University Press.

Green, M. C. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes*, *38*(2), 247–266. DOI: https://doi.org/10.1207/s15326950dp3802_5

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*(5), 701–721. DOI: https://doi.org/10.1037/0022-3514.79.5.701

Green, M. C., & Brock, T. C. (2002). In the mind's eye: Transportation-imagery model of narrative persuasion. In M. C. Green, J. J. Strange & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 315–341). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers. DOI: https://doi.org/10.1533/9780857095657.2.362

Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, *14*(4), 311–327. DOI: https://doi.org/10.1111/j.1468-2885.2004.tb00317.x

Green, M. C., Kass, S., Carrey, J., Herzig, B., Feeney, R., & Sabini, J. (2008). Transportation across media: Repeated exposure to print and film. *Media Psychology*, *11*(4), 512–539. DOI: https://doi.org/10.1080/15213260802492000

Hall, A., & Zwarun, L. (2012). Challenging entertainment: Enjoyment, transportation, and Need for Cognition in relation to fictional films viewed online. *Mass*

Communication and Society, *15*(3), 384–406. DOI: https://doi.org/10.1080/15205436.2011.583544

Hartung, F., Burke, M., Hagoort, P., & **Willems, R. M.** (2016). Taking perspective: Personal pronouns affect experiential aspects of literary reading. *PLoS ONE*, *11*(5), 1–18. DOI: https://doi.org/10.1371/journal.pone.0154732

Hartung, F., Withers, P., Hagoort, P., & **Willems, R. M.** (2017). When fiction is just as real as fact: No differences in reading behavior between stories believed to be based on true or fictional events. *Frontiers in Psychology*, *8*, 1–14. DOI: https://doi.org/10.3389/fpsyg.2017.01618

Jacobs, A. M. (2015). Neurocognitive poetics: Methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, *9*, 1–22. DOI: https://doi.org/10.3389/fnhum.2015.00186

Jacobs, A. M., & **Willems, R. M.** (2018). The Fictive Brain: Neurocognitive Correlates of Engagement in Literature. *Review of General Psychology*, *22*(2), 147–160. DOI: https://doi.org/10.1037/gpr0000106

Janssen, T., Braaksma, M., & **Couzijn, M.** (2009). Self-questioning in the literature classroom: Effects on students' interpretation and appreciation of short stories. *L1 Educational Studies in Language and Literature*, *9*(1), 91–116. DOI: https://doi.org/10.17239/L1ESLL-2009.09.01.05

Johnson, D. R., Cushman, G. K., Borden, L. A., & **McCune, M. S.** (2013). Potentiating empathic growth: Generating imagery while reading fiction increases empathy and prosocial behavior. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(3), 306–312. DOI: https://doi.org/10.1037/a0033261

Knoop, C. A., Wagner, V., Jacobsen, T., & **Menninghaus, W.** (2016). Mapping the aesthetic space of literature "from below." *Poetics*, *56*, 35–49. DOI: https://doi.org/10.1016/j.poetic.2016.02.001

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

Kraaykamp, G., & **van Eijck, K.** (2005). Personality, media preferences, and cultural participation. *Personality and Individual Differences*, *38*, 1675–1688. DOI: https://doi.org/10.1016/j.paid.2004.11.002

Kuijpers, M., Douglas, S., & **Kuiken, D.** (2018). Personality traits and reading habits that predict absorbed narrative fiction reading. *Psychology of Aesthetics, Creativity, and the Arts*, 1–15. DOI: https://doi.org/10.1037/aca0000168

Kuijpers, M. M., Hakemulder, F., Tan, E. S., & **Doicaru, M. M.** (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. DOI: https://doi.org/10.1075/ssol.4.1.05kui

Kuiken, D., & **Douglas, S.** (2017). Forms of absorption that facilitate the aesthetic and explanatory effects of literary reading. In F. Hakemulder, M. M. Kuijpers, E. S. Tan, K. Bálint & M. M. Doicaru (Eds.), *Narrative Absorption* (pp. 217–249). Amsterdam/Philadelphia: John Benjamins Publishing Company. DOI: https://doi.org/10.1075/lal.27.12kui

Kuzmičová, A. (2012). Presence in the reading of literary narrative: A case for motor enactment. *Semiotica*, *189*(1), 23–48. DOI: https://doi.org/10.1515/semi.2011.071

Kuzmičová, A. (2014). Literary narrative and mental imagery: a view from embodied cognition. *Style*, *48*(3), 275–293. DOI: https://doi.org/10.5325/style.48.3.275

Kuznetsova, A., Brockhoff, P. B., & **Christensen, R. H. B.** (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. DOI: https://doi.org/10.18637/jss.v082.i13

Lao, C. Y., & **Krashen, S.** (2000). The impact of popular literature study on literacy development in EFL: more evidence for the power of reading. *System*, *28*(2), 261–270. DOI: https://doi.org/10.1016/S0346-251X(00)00011-7

Lee, J., Schallert, D. L., & **Kim, E.** (2015). Effects of extensive reading and translation activities on grammar knowledge and attitudes for EFL adolescents. *System*, *52*, 38–50. DOI: https://doi.org/10.1016/j.system.2015.04.016

Lenth, R. (2018). emmeans: Estimated Marginal Means, aka Least-Squares Means. *R package version 1.2.4.* Retrieved from https://cran.r-project.org/package=emmeans

Mak, M., & **Willems, R. M.** (2019). Mental simulation during literary reading: Individual differences revealed with eye-tracking. *Language, Cognition and Neuroscience*, *34*(4), 511–535. DOI: https://doi.org/10.1080/23273798.2018.1552007

Malanchini, M., Wang, Z., Voronin, I., Schenker, V. J., Plomin, R., Petrill, S. A., & **Kovas, Y.** (2017). Reading self-perceived ability, enjoyment and achievement: A genetically informative study of their reciprocal links over time. *Developmental Psychology*, *53*(4), 698–712. DOI: https://doi.org/10.1037/dev0000209

Mar, R. A., & **Oatley, K.** (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *3*(3), 173–192. DOI: https://doi.org/10.1111/j.1745-6924.2008.00073.x

Miall, D. S., & **Kuiken, D.** (1995). Aspects of literary response: A new questionnaire. *Research in the Teaching of English*, *29*(1), 37–58. DOI: https://doi.org/10.2307/40171422

Mol, S. E., & **Jolles, J.** (2014). Reading enjoyment amongst non-leisure readers can affect achievement in secondary school. *Frontiers in Psychology*, *5*, 1–10. DOI: https://doi.org/10.3389/fpsyg.2014.01214

Nijhof, A. D., & **Willems, R. M.** (2015). Simulating fiction: Individual differences in literature comprehension revealed with fMRI. *PLoS ONE*, *10*(2), 1–17. DOI: https://doi.org/10.1371/journal.pone.0116492

Oatley, K. (2016). Fiction: Simulation of social worlds. *Trends in Cognitive Sciences, 20*(8), 618–628. DOI: https://doi.org/10.1016/j.tics.2016.06.002

Panero, M. E., Weisberg, D. S., Black, J., Goldstein, T. R., Barnes, J. L., Brownell, H., & **Winner, E.** (2016). Does reading a single passage of literary fiction really improve theory of mind? An attempt at replication. *Journal of Personality and Social Psychology, 111*(5), e46–e54. DOI: https://doi.org/10.1037/pspa0000064

Peters, G. J. Y. (2014). The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist, 16*(2), 56–69.

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Retelsdorf, J., Köller, O., & **Möller, J.** (2011). On the effects of motivation on reading performance growth in secondary school. *Learning and Instruction, 21*, 550–559. DOI: https://doi.org/10.1016/j.learninstruc.2010.11.001

Revelle, W. (2014). Classical Test Theory and the Measurement of Reliability. In *Psychometric Theory* (pp. 205–239). Retrieved from http://www.personality-project.org/r/book/

Ryan, M. L. (2001). *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media.* Baltimore/London: Johns Hopkins University Press.

Samur, D., Tops, M., & **Koole, S. L.** (2017). Does a single session of reading literary fiction prime enhanced mentalising performance? Four replication experiments of Kidd and Castano (2013). *Cognition and Emotion*, 1–15. DOI: https://doi.org/10.1080/02699931.2017.1279591

Schutte, N. S., & **Malouff, J. M.** (2004). University student reading preferences in relation to the big five personality dimensions. *Reading Psychology, 25*, 273–295. DOI: https://doi.org/10.1080/02702710490522630

Stanovich, K. E., & **West, R. F.** (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*(4), 402–433. DOI: https://doi.org/10.2307/747605

Tukachinsky, R. (2014). Experimental manipulation of psychological involvement with media. *Communication Methods and Measures, 8*(1), 1–33. DOI: https://doi.org/10.1080/19312458.2013.873777

Van Dis, A. (1986). *Casablanca.* Amsterdam, The Netherlands: Meulenhoff.

Van Essen, R. (2014). De mensen die alles lieten bezorgen. In *Hier wonen ook mensen* (pp. 113–123). Utrecht: Atlas Contact.

Weibel, D., Wissmath, B., & **Mast, F. W.** (2011). Influence of mental imagery on spatial presence and enjoyment assessed in different types of media. *Cyberpsychology, Behavior, and Social Networking, 14*(10), 607–612. DOI: https://doi.org/10.1089/cyber.2010.0287

Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System, 36*(4), 661–672. DOI: https://doi.org/10.1016/j.system.2008.04.003