Educational Psychology
An International Journal of Experimental Educational Psychology
ISSN: 0144-3410 (Print) 1469-5820 (Online) Journal homepage: https://www.tandfonline.com/loi/cedp20

Inspecting a picture before reading affects attentional processing but not comprehension

Stephanie I. Wassenburg, Björn B. de Koning, Lisanne T. Bos & Menno van der Schoot

To cite this article: Stephanie I. Wassenburg, Björn B. de Koning, Lisanne T. Bos & Menno van der Schoot (2019): Inspecting a picture before reading affects attentional processing but not comprehension, Educational Psychology, DOI: 10.1080/01443410.2019.1645306

To link to this article: https://doi.org/10.1080/01443410.2019.1645306

Published online: 04 Aug 2019.

EDUCATIONAL PSYCHOLOGY
https://doi.org/10.1080/01443410.2019.1645306

Inspecting a picture before reading affects attentional processing but not comprehension

Stephanie I. Wassenburg[a,b], Björn B. de Koning[a], Lisanne T. Bos[c] and Menno van der Schoot[c]

[a]Department of Psychology, Education, and Child Studies, Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands; [b]Human-Computer Interaction Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA; [c]Department of Educational and Family Studies and LEARN!, Faculty of Behavioural and Movement Sciences, Research Institute for Learning and Education, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

ABSTRACT

This study investigated whether presenting a picture before reading can encourage situation-model construction. We compared two conditions ($n = 30$) which differed in whether a picture of the initial situation described in a narrative text was presented before reading (i.e. pictorial-support condition) or not (i.e. no-picture condition). Situation-model construction was measured using both process- and product-oriented measures. Eye-tracking data indicated online resource allocation to the different levels of text representation: surface, textbase, and situation model. Literal text questions and inference questions were used as an offline indication of textbase and situation-model processing, respectively. The results showed that a picture presented before reading led to a redistribution of processing resources during reading, evidenced by a shift from textbase to situation-model processing. This attentional shift did not translate into higher comprehension scores. The results were interpreted in line with multimedia learning theories suggesting pictures can serve as a mental scaffold for situation-model construction.

ARTICLE HISTORY
Received 26 October 2018
Accepted 14 July 2019

KEYWORDS
Reading comprehension; pictorial support; situation model; resource allocation; eye-tracking

Introduction

It is generally acknowledged that a text can be mentally represented at different levels, rather than reading comprehension being an all-or-nothing process. Three levels of representation – surface, textbase, and situation model – have been distinguished by Kintsch (1998). On the surface level, readers construct a mental representation of the literal text, which features, for example, the exact words, clauses, and their syntactic relations (McNamara & Magliano, 2009). On the textbase level, the representation

CONTACT Menno van der Schoot m.vander.schoot@vu.nl Department of Educational and Family Studies and LEARN!, Research Institute for Learning and Education, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

features more meaningful aspects of the text, such as meanings of words, clauses, and sentences, as well as their mutual relations, are represented (Kintsch & Rawson, 2005). Finally, the situation-model level exceeds the literal text information that is represented by the former two levels by featuring a coherent and integrated non-verbal representation of the situation described in a text (Kintsch, 1998; Zwaan & Radvansky, 1998). Importantly, a situation model is constructed by integrating and enriching the literal content of the text with the readers' prior knowledge (Stine-Morrow, Gagne, Morrow, & DeWall, 2004; Wiley, Griffin, & Thiede, 2005). This way, the situation model becomes a coherent and richly connected mental representation of the story content, comparable to a representation of an actually experienced situation (e.g. Fletscher, 1994). Hence, the construction of a situation model, in particular, is associated with deep comprehension (van der Schoot, Horsley, & van Lieshout, 2010). Nevertheless, readers do not always construct situation models from a text, either because they lack knowledge of appropriate reading comprehension strategies (e.g. Cain & Oakhill, 1999; Elbro & Buch-Iversen, 2013; Wassenburg, Bos, de Koning, & van der Schoot, 2015) or because instruction does not encourage them to do so (e.g. van der Schoot et al., 2010). The present study addresses the latter issue by investigating how instruction can increase readers' effort to construct a situation-model representation.

Construction of a situation model occurs gradually by monitoring different situational dimensions of the narrative text such as space, time, protagonist, causation, and intentionality (Wassenburg, Beker, van den Broek, & van der Schoot, 2015; Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995). Readers often do this by mentally situating themselves within a narrative text scene (Zwaan, 1999). From that perspective, readers move along with the main protagonist and experience the unfolding actions and developments in the story as the protagonist would do. This helps them to keep track of the various situational dimensions and keep their situation model up-to-date with the narrative text (e.g. Rinck & Bower, 1995). Narrative texts, however, often contain discontinuities in space and time, for example, when the protagonist suddenly is in a different setting or when a week has passed. When this happens, especially when shifts are so large that they must be considered part of a new situation, updating the situation model is more difficult and comprehension is impeded (Zwaan & Radvansky, 1998). To resolve the breaks in continuity, readers should infer missing information based on real-world knowledge (e.g. van den Broek, 1997). Spatial or temporal discontinuities require, therefore, extra effort and processing time (van den Broek & Lorch, 1993; Zwaan, 1996; Zwaan, Magliano, et al., 1995).

Because the additional effort invested in constructing a situation model is related to deeper text processing, it is important that instruction encourages readers to engage in these demanding processes. In this context, research has shown that instructions to recall text lead to allocation of processing resources towards the surface and textbase representation (Stine-Morrow, Milinder, Pullara, & Herman, 2001), whereas specific situation-model instructions can lead to a shift in processing effort from textbase towards a situation-model representation (Millis, Simon, & Tenbroek, 1998; Stine-Morrow et al., 2004; van der Schoot et al., 2010; Zwaan, Magliano, et al., 1995). To investigate the effect of situation-model instruction on both online and offline reading processes, van der Schoot et al. (2010) combined eye-tracking data and reading comprehension outcomes. More specifically, in a situational-instruction condition, they explicitly asked participants to imagine, as vividly as possible, the events and situations that were described in the text. As an example, it was explained that they could adopt the perspective of the protagonist. In the control condition, readers were only instructed to read the text for comprehension. Providing verbal instruction designed to encourage situation-model construction led to the redistribution of attentional resources from textbase to situation-model processing and resulted in better comprehension. The effects, however, were small and only present for poor comprehenders, which fits the broader literature on the limited effects of an 'imagination instruction' on deeper processing of text (de Koning & van der Schoot, 2013). Furthermore, and of more relevance to the present study, most of the studies on the effects of instruction on situation-model construction – including the one by van der Schoot et al. (2010) – have focused on manipulating verbal instruction (e.g. Andreassen & Bråten, 2011; Guthrie et al., 2009; Houtveen & van de Grift, 2007; Stine-Morrow et al., 2001). However, based on the reasoning that pictures – like situation models – are non-verbal representations of the events described in a text, greater benefits may be obtained by providing pictorial support.

Pictorial support

Pictures have long been known to support comprehension (e.g. Carney & Levin, 2002; Eitel & Scheiter, 2015; Glenberg & Langston, 1992; Newton, 1994; Schnotz & Bannert, 2003). Most research investigating the supportive function of pictures for processing text is performed in the field of multimedia learning. According to theories of multimedia learning (e.g. Mayer, 2009; Schnotz, 2002), surface representations – of which the modality is similar to that of the presented materials – are processed first. This means that if a picture is presented, the learner creates a perceptual representation of the picture onto which the text's semantic representations can be mapped. In this sense, a picture provides the 'blueprint' of a situation model (Nyhout & O'Neill, 2017) and hence may serve as a mental scaffold to facilitate its construction.

In multimedia learning studies, text and pictures are typically presented simultaneously (e.g. Mayer, 2009; Schnotz & Bannert, 2003; Schüler, Arndt, & Scheiter, 2015). It has been shown, for example, that providing a picture, which depicts the initial state of the described situation, along with the first paragraph of a narrative text helps children to construct an appropriate situation model (Newton, 1994). Additional evidence for the supportive function of pictures for reading comprehension comes from a study by Pike, Barnes, and Barron (2010) who showed that a picture presented directly below the text improved inference generation. There is also some initial evidence that sequencing picture and text support comprehension (Eitel & Scheiter, 2015). In particular, Eitel, Scheiter, Schüler, Nyström, and Holmqvist (2013) showed that presenting a picture before a reading-related task for a group of adult readers resulted in better comprehension of an expository text when compared to a group who did not see a picture before the task. Together, these results suggest that pictures facilitate the construction of a situation model. These studies, however, have focused on expository texts and have only measured the product of situation-model processing, by using recall, interviews, or comprehension questions.

Expository texts usually pose a larger challenge for readers than narrative texts, because they present more new and abstract concepts, and have a less predictable structure (Best, Floyd, & McNamara, 2008; Kraal, Koornneef, Saab, & van den Broek, 2018). Due to such differences, narrative and expository texts also elicit different processing strategies (Kraal et al., 2018). Best et al. (2008) showed, for example, that comprehension of expository text was more dependent than narrative texts on world knowledge. Furthermore, expository texts elicit processing of literal text and details, whereas narrative texts evoke processing of themes and structure (Wolfe, 2005). Despite these differences, it is expected that pictorial support may enhance situation-model construction in both types of texts, as the construction of a situation model is necessary for comprehension and is based on inferences from the explicit text, the reader's prior knowledge, and their interactions (Graesser, Singer, & Trabasso, 1994). It is important, however, that these differences are taken into consideration while developing pictorial support. A situation model of an expository text refers to how well the reader has integrated the textbase information of a certain subject matter with their prior knowledge on that matter (Best et al., 2008). Therefore, readers of expository texts may benefit from pictures providing the information they do not possess yet or complex relations. A situation model of a narrative text refers to how well the reader understands the information about the characters, actions, settings, and events. Hence, readers of narrative texts may benefit mostly from pictures containing information with regard to protagonists or narrative events which readers can use to activate the correct event schema.

Therefore, it is relevant to consider text type when investigating the effects of presenting a picture before the text. The present study extends previous work in this area in two ways. First, we will study the effect of showing a picture before reading using a narrative text. Second, we will combine product-oriented and process-oriented measures to get insight into online 'picture-enhanced' situation-model building and how this contributes to post-reading comprehension. Textbase and situation-model understanding of the text will be measured with literal and comprehension questions, respectively, while distribution of attentional resources over the different levels of text representation (i.e. surface, textbase, situation model) will be measured with an eye tracker using the resource allocation approach.

Resource allocation approach

The resource allocation approach focuses on how readers distribute their available mental resources over all the simultaneous demands that a text makes upon them, rather than on how well they execute specific component processes in reading (van der Schoot et al., 2010). In this respect, the resource allocation approach deviates from standard classical eye-movement techniques measuring perceptual, visuomotor, and cognitive-linguistic processes involved in reading (e.g. Rayner, 1998). The approach assumes that self-regulated allocation of time to reading reflects attention to or effort put into the processing of different features of the text with longer reading times indicating a more thorough analysis of textual features (Stine-Morrow et al., 2004). Accordingly, if reading times are analysed as a function of specific text variables that are associated with the processing of the different levels of text representation (i.e. surface, textbase, and situation model), the results provide an indication of the effort that readers put into processing those respective levels of text representation. In other words, regressing reading times of sentences on text variables can provide information on the extent to which readers allocate process resources to the different representations of a text (e.g. Aaronson & Scarborough, 1977; Graesser, Hoffman, & Clark, 1980; Millis et al., 1998; Stine, 1990; Stine-Morrow et al., 2001; van der Schoot et al., 2010; Zwaan, Magliano, et al., 1995).

Specifically, a surface representation of the literal text, featuring the exact words, clauses, and their syntactic relations, is associated with allocating attention to and processing of the orthographic content of a text (e.g. number of syllables, word frequency). Textbase representations (i.e. representing the meanings of words and their mutual relations) require, for example, the processing of propositions and new concepts. Constructing a situation-model representation involves monitoring causal, temporal, and spatial information in the text (Radvansky, Zwaan, Curiel, & Copeland, 2001; Stine-Morrow et al., 2001, 2004; van der Schoot et al., 2010; Zwaan, Magliano, et al., 1995). Following this reasoning, a passage introducing new concepts requires additional processing effort on the textbase level, whereas temporal shifts require additional processing efforts on the situation-model level. Both text variables may slow down reading, but only if the reader allocates attention and effort to these aspects. Stated differently, if readers mostly focus on the surface and textbase representation of the text, text variables such as the number of syllables and number of new concepts will have a greater impact on reading times than spatial shifts. As mentioned before, the situation model level is especially important for deep text comprehension (e.g. van der Schoot et al., 2010).

Present study

In the present study, participants read a narrative text rich in spatial shifts. It is assumed that readers must build a new situation model, resulting in longer reading times, for every spatial shift that is encountered (Zwaan & Radvansky, 1998). Hence, situation-model processing is important for accurate text comprehension. The text used in the present study was preceded by a picture (see Figure 1) for half of the participants before they started reading. The picture illustrated the begin state of the narrative text and was designed to encourage (1) processing of spatial information by showing spatial relationships between the depicted entities (e.g. the mouse is depicted mid-jump onto a train, and the lines behind him suggest motion and direction) and (2) generation of inferences by providing cues that can be used to integrate prior knowledge (e.g. waving people suggest the train is leaving soon, and the mouse carries a bindle suggesting he will be traveling for a longer period of time). In other words, the picture was intended to help readers to construct a situation model of the initial situation of the story in particular. If participants use this picture as a primary building block for the construction of a situation model of the described situation (cf.

Figure 1. The cover of the narrative titled 'The journey of the mouse'.

Eitel et al., 2013), the picture should function as a visuospatial scaffold for (further) situation-model processing. Accordingly, readers can begin with and maintain a more effective situation model of the narrative text as it progresses than someone who has not seen the picture (Eitel & Scheiter, 2015). Picture-enhanced situation-model construction would be evidenced by a redistribution of allocated time from mainly text-base levels towards situation-model processing, and a corresponding increase in reading comprehension outcomes on questions tapping into the situation-model representation. Such a shift in processing and increased comprehension outcomes (at the situation-model level) was not expected in a no-picture condition in which participants read the narrative text without receiving a picture beforehand.

Method

Participants

Adults with neurological disorders such as attention deficit hyperactivity disorder, autism, and dyslexia were not allowed to participate. All participants were native Dutch speakers with normal or corrected-to-normal vision. From the initial sample of 36 participants, data of six participants were excluded from analyses due to problems with their eye-tracking recording. The final sample consisted of 30 adult participants (21 women) with a mean age of 28.12 years ($SD = 10.73$ years). Participants were randomly assigned to either the pictorial-support condition ($n = 16$) or the no-picture condition ($n = 14$). None of the participants had read or heard about the story before.

Materials

Text

The narrative text (see Appendix) was based on a passage from the book '*Het Sleutelkruid'* by Paul Biegel (published in English as 'The King of The Copper Mountains') and was about a mouse on a journey. The text contained many spatial shifts and had a clear protagonist (the mouse) and an episodic structure. The text was three pages long and consisted of 30 sentences (10 sentences per page) and 404 words (the number of words per sentence: $M = 13.77$, $SD = 4.46$).

Cover

The cover, presented in Figure 1, illustrated the beginning of the story where the mouse jumped onto the train to start his journey (i.e. first spatial discontinuity). It was designed and drawn for this study by a professional draftsman. This cover was only presented in the pictorial-support condition.

Comprehension

To measure comprehension of the story, eight multiple-choice questions were presented (e.g. 'The mouse fell of the car because …' [a] the car was driving so fast, [b] there was heavy rainfall, [c] the car got an accident, or [d] the roof got slippery). The questions differentiated between the textbase and situation-model level of the story. Four questions focused on the literal text (textbase level), whereas the other four questions required making inferences beyond the literal text (situation-model level). For literal questions and inference questions, separate reading comprehension scores were calculated as the proportion of correct answers.

Text variables

To investigate readers' relative allocation of attentional resources to surface, textbase, and situation-model processing, each sentence of the experimental text was initially coded on a number of associated text variables (see Table 1). To avoid multicollinearity in later regression models, correlational analyses were performed to determine whether certain text variables should be excluded. For surface processing, sentences were coded on a number of letters, syllables, and words (orthographic decoding; Stine-Morrow et al., 2001), and word frequency (based on the SUBTLEX-NL database; Keuleers, Brysbaert, & New, 2010). A log transformation was performed on mean word frequencies to meet the assumption of normal distributions. A number of letters, syllables, and words were highly correlated ($r > 0.90$). Therefore, a number of letters and number of words were excluded, and a number of syllables and average log word

Table 1. Descriptive statistics for the text variables per sentence.

	M	*SD*
Total number of syllables (range: 5–24)	16.37	4.41
Number of new concepts (range: 0–5)	2.21	1.40
Spatial discontinuities (range: 0–2)	0.76	0.73
Sentence position (range 1–10)	5.65	2.80
Mean (log) word frequency (range 4–5)	4.64	0.25

frequency were included as predictors of surface processing (van der Schoot et al., 2010). For textbase processing, sentences were coded on a total number of concepts and number of new concepts (Graesser et al., 1980; Millis et al., 1998). Because of a high correlation ($r = 0.68$), only the number of new concepts was included for further analyses as new concepts require more processing effort than old concepts (Stine-Morrow et al., 2004). For situation-model processing, sentences were coded on temporal and spatial discontinuity (Radvansky et al., 2001). For example, a spatial discontinuity was present when the protagonist had moved to a different region. A region is a segment of space with clearly defined boundaries (e.g. a shift from a train to a boat). The number of temporal and spatial shifts per sentence were not correlated ($r = 0.03$), but as the number of temporal shifts was not related to reading times ($r = 0.03$), only spatial shift was included for further analyses. Additionally, sentence position (1 = top of the page, 10 = bottom of the page) was added as a control variable as this may influence reading times.

Eye-fixation data

Eye-movement data were recorded using the Tobii TX300 eye tracker (Tobii, Stockholm, Sweden), sampling at 300 Hz from both eyes. This 23″ 1920 × 1080 pixels widescreen monitor contains an eye-tracking camera and infrared LEDs mounted on the monitor bezel. During the reading of the text, raw gaze data were recorded and transformed into fixations using the Tobii Fixation Filter (velocity threshold of 35 pixels/window). The Tobii Fixation Filter is an implementation of a classification algorithm proposed by Olsson (2007) and is currently used as the default fixation algorithm in Tobii's analysis software. Each sentence of the text – presented without line breaks – was defined as an area of interest (AOI). Because the resource allocation methodology requires sentence reading times (i.e. total time spent on reading a sentence), the Total Visit Duration (TVD) was calculated for each of the 30 AOIs (i.e. sentences), based on the fixations identified by the fixation filter. The TVD is a metric that is used to indicate the total time that a reader has spent fixating on an AOI. It contains all visits and revisits and is calculated as the sum of all visit durations (Holmqvist et al., 2011). The fixation on the title AOI was removed for further analyses (only the pictorial-support group saw the title of the story on the cover, and both groups saw it at the start of the text). Tobii Studio 3.3.0 software was used for text presentation, eye calibration, data recording, and analyses of fixation data.

Procedure

Participants were tested in a university laboratory. They were seated with approximately 70 cm distance between their eyes and the computer screen. At the start of the experiment, the eye tracker was adjusted and calibrated using a five-point calibration grid presented on the screen. If calibration was non-optimal, the process was repeated. After calibration, participants read instructions stating that they were supposed to sit as still as possible and to use the spacebar to move on to the next screen. They were further instructed to silently read a text on which they would be tested for their comprehension afterward. Only in the pictorial-support condition, participants were additionally instructed to look at the presented cover (presented for seven seconds) before the text was presented and use this to help them comprehend the text. In the no-picture condition, the text followed the instructions immediately. The text was divided over three pages. Each page (containing 10 sentences) was presented statically on the computer screen in order to measure eye fixations on the different sentences (i.e. TVDs). Participants went to the next page by pressing the space bar; once on the next page, it was not possible to go back to a previous page. Finally, participants answered the eight comprehension questions that were presented on the screen by typing their answer in a text box. During the experiment, the researcher was always present in the room. The experiment lasted approximately 20 minutes.

Resource allocation analysis

First, two overall regression analyses were conducted to investigate to what extent the chosen text variables affect TVD. This was done separately for the pictorial-support and no-picture condition, as the TVD may be differentially influenced in the two conditions. Second, individual regression analyses were conducted for every participant separately to investigate how text variables affect their individual TVD, and subsequently to create three new variables consisting of the resulting standardised regression coefficients – one for each representational level. In other words, resource allocation to surface, textbase, and situation-model processing is operationalised as changes in reading time as a function of the associated text variables. Finally, these standardised regression coefficients were used as a dependent variable to conduct a mixed analysis of variance to investigate the effect of pictorial support on resource allocation.

Results

Resource allocation

Overall effects of text variables on TVD

Two regression analyses (for the two conditions separately) were run with TVD per participant per sentence as dependent variable and the text variables syllables, average log word frequency, new concepts, spatial discontinuities, and sentence position as predictors. All variables were significantly correlated with TVD. Correlations between text variables ranged from $r < 0.001$ to $r = 0.42$, which did not cause multicollinearity problems. Table 2 shows the standardised regression coefficients separately for the no-picture and pictorial-support conditions. Together, the five text variables significantly predicted variance in TVDs in both the no-picture condition ($R^2 = 0.22$, $p < .001$) and the pictorial-support condition ($R^2 = 0.19$, $p < .001$). As can be seen in Table 2, however, the relative contributions of the text variables differed per condition. Although participants in both conditions allocated a similar amount of resources to text decoding (i.e. the number of syllables), the two conditions differed in the extent to which resources were allocated to the textbase and situation-model representations. The number of new concepts was a stronger predictor of TVD in the no-picture

Table 2. Standardised regression coefficients for the text variables per condition.

	No picture		Pictorial support	
	β	*t*	*β*	*t*
Syllables	0.30	5.91**	0.25	5.15**
New concepts	0.21	3.70**	0.14	2.73*
Spatial discontinuities	0.05	1.02	0.16	3.35**
Sentence position	0.02	0.45	0.04	0.95
Mean word frequency	−0.08	−1.78	−0.04	−0.77

*$p < .01$,
**$p < .001$.

condition than in the pictorial-support condition. Moreover, the number of spatial discontinuities was a significant predictor in the pictorial-support condition, but not in the no-picture condition. Average log word frequency and sentence position did not predict a significant amount of variance in TVDs and were therefore removed from further analyses.

Relative resource allocation per participant

Individual regression analyses (i.e. for every participant separately) were performed with TVD per sentence as dependent variable and the three text variables syllables, new concepts, and spatial discontinuities as predictors (see also Radvansky et al., 2001; Stine-Morrow et al., 2001, 2004; van der Schoot et al., 2010; Zwaan, Magliano, et al., 1995). The output provided an individual regression coefficient for each representation level for each participant, which reflects the relative resource allocation to surface, textbase, and situation-model level processing. Previous research has indicated that such coefficient patterns show significant inter-individual variability and reliability over time (Stine-Morrow et al., 2001). Including the number of syllables in the regression analyses was necessary to control for individual differences in surface-level reading and calculate the unique contributions of textbase and situation-model processing over and above surface-level reading. However, no differences were expected in surface-level reading times between the two conditions because it is independent of comprehension (i.e. processing more syllables requires more time simply because the sentence is longer). Therefore, the number of syllables was excluded from further analyses that focus on differences in resource allocation to lower (i.e. textbase) and higher-level (i.e. situation model) sense-making processes.

Effect of pictorial support

To investigate the differences in contributions of the two highest levels (i.e. new concepts and spatial discontinuities) to TDVs between conditions, we conducted a 2×2 mixed analysis of variance (ANOVA) on the standardised beta coefficients with Representation level as within-subject factor (textbase vs. situation model) and Picture (pictorial support vs. no picture) as between-subject variable. The significant main effect of Representation level, $F(1, 28) = 5.03$, $p = .033$, $\eta_p^2 = 0.15$, indicates that participants generally allocated more resources to textbase processes (i.e. new concepts; $M = .23$, $SD = .21$) than to situation-model processes (i.e. spatial discontinuities;

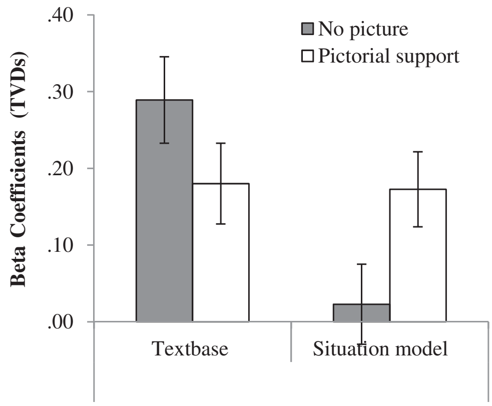

Figure 2. Resource allocation (error bars depict standard errors from the mean) as a function of Representation level and Picture.

$M = 0.10$, $SD = 0.21$). Overall beta coefficients (across Representation level) did not differ as a function of Picture ($F < 1$). More importantly, however, there was a significant interaction between Representation level and Picture, $F(1, 28) = 4.50$, $p = .043$, $\eta_p^2 = 0.14$. As can be seen in Figure 2, in the pictorial-support condition there was a shift in attentional processing such that more resources were allocated to processing situation-model features in the text and less to processing textbase features as compared to the no-picture condition. Follow-up tests confirmed that the pictorial-support group allocated significantly more resources to the situation-model level than the no-picture group, $t(28) = -2.10$, $p = .045$, Cohen's $d = 0.80$. Within the pictorial-support condition, the amount of processing resources allocated to situation-model features ($M = 0.17$, $SD = 0.19$) was comparable to that of textbase features ($M = 0.18$, $SD = 0.17$), $t(15) = 0.10$, $p = .923$, Cohen's $d = 0.06$, indicating that processing resources were approximately equally distributed over textbase and situation-model features in the text after having inspected a picture before reading. The no-picture group instead allocated significantly more resources to textbase processing ($M = 0.29$, $SD = 0.25$) than to situation-model processing ($M = 0.02$, $SD = 0.19$), $t(13) = 2.71$, $p = .018$, Cohen's $d = 1.22$. The allocation of resources to textbase processing was statistically similar across the two conditions, $t(28) = 1.41$, $p = .169$, Cohen's $d = 0.53$.

Comprehension questions

The overall proportion correct for the comprehension questions was 0.69 ($SD = 0.15$). The accuracy scores on the comprehension questions (shown in Figure 3) were submitted to a 2×2 mixed ANOVA with Representation level (textbase vs. situation model) as within-subject factor and Picture (pictorial support vs. no picture) as between-subject variable. As can be seen in Figure 3, comprehension was higher for literal (i.e. textbase level) questions ($M = 0.78$, $SD = .21$) than for inference (i.e. situation-model level) questions ($M = 0.61$, $SD = 0.20$), which was reflected by a significant main effect of Representation level, $F(1, 28) = 9.71$, $p = .004$, $\eta_p^2 = 0.26$. More importantly, the results showed that providing a picture did not affect the proportion of correct answers either for literal or for inference questions, as was evidenced by the absence of a significant main effect of Picture, $F(1, 28) = 1.13$, $p = .297$, $\eta_p^2 = 0.04$, and

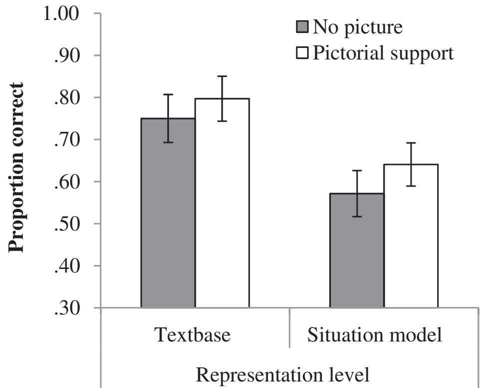

Figure 3. Comprehension scores (error bars depict standard errors from the mean) as a function of Representation Level and Picture.

Picture by Representation level interaction, $F(1, 28) = 0.04$, $p = .837$, $\eta_p^2 < 0.01$. This indicates that the shift in online resource allocation due to a picture (i.e. from textbase to situation-model level) did not translate into improved inference-related comprehension scores.

Discussion

The present study investigated the effects of pictorial support before reading a narrative text on situation-model processing. This was examined on both process-oriented and product-oriented measures. Reading times were calculated using eye-tracking and analysed as a function of specific text variables associated with the online surface, textbase, and situation-model processing. Reading comprehension questions on both textbase and situation-model level were used as product measures.

The results showed that participants who were instructed to look at a picture depicting the start situation of a story before reading allocated significantly more resources to situation-model processes than participants who did not receive such pictorial support. As expected, presenting a picture before reading led to a redistribution of online processing resources, causing a shift in attention from the textbase to a situation-model representation. This is in line with expectations based on previous research showing that instruction can influence online resource allocation (e.g. Millis et al., 1998; Stine-Morrow et al., 2004; van der Schoot et al., 2010; Zwaan, Magliano, et al., 1995) and that providing pictures before or during reading can support inference generation, recall, and reading comprehension for expository texts (e.g. Eitel et al., 2013; Eitel & Scheiter, 2015; Newton, 1994; Pike et al., 2010). This study extends these results by showing that briefly looking at a picture pre-reading can cause a redistribution of online processing resources in a narrative text. Moreover, the present findings show that situation-model-focused instructions not only influence attentional processing in children who are poor comprehenders (cf. van der Schoot et al., 2010) but also work for an adult sample consisting of readers who are assumed to be skilled in reading.

Contrary to expectations, the attentional shift to situation-model processing due to the offered pictorial support did not result in corresponding improvements in reading comprehension scores. Irrespective of whether or not readers looked at a picture before reading, their accuracy on both textbase and situation-model level comprehension questions was comparable. These findings align with other studies outside of a reading comprehension context where the disconnect between online attentional processing and cognitive outcomes is reported. In the field of learning from instructional visualisations for example, de Koning, Tabbers, Rikers, and Paas (2010) showed that attentional cues designed to guide attention to task-relevant information in the visualisation helped learners to focus their attention on the relevant content, but this did not lead to better learning outcomes on a subsequent test than when having studied the visualisation without attention-directing cues. Similarly, in the field of multimedia learning, Pouw, Rop, de Koning, and Paas (2019) observed less integrative eye movements during learning of text-picture instructions when the spatial distance between text and picture was increased even though learning outcomes remained similar in the large and small spatial-distance conditions. The results of this study demonstrate that such effects also exist when processing narrative text after having inspected a picture. However, this conclusion needs to be interpreted with caution given that the present comprehension measure contained a relatively small number of items, the study involved a limited number of participants, and we used only one text (with the picture presented before it) in our experiment. This may have reduced the chance to detect an effect on comprehension. Furthermore, even though the narrative (appropriate for young adolescents) should be relatively easy to read for adults, due to the between-subject design we cannot exclude the possibility that individual differences in reading and comprehension skills have impacted our results.

The fact that we found a redistribution of resources in online processing but no differences in outcome measures could indicate that the manipulation (showing a picture) provided too subtle guidance and did not explicitly encourage readers to process the text at an appropriate (deeper) level that is necessary to perform better on the inference-related questions. Alternatively, it is possible that even with minimal processing effort directed to a situation-model representation, participants could perform relatively well on the comprehension test, for example because the difficulty level of the comprehension questions was rather low for this participant population (i.e. adults reading a narrative on adolescents' level). In fact, in both conditions participants provided a suitable answer on at least 50% of the comprehension questions. Another possibility is that the comprehension test provided a learning moment in and of itself where learners engaged in deeper processing of the just-read text *after* instead of *during* reading. That is, by reading the questions, the participants may have been challenged to deeply process the text and draw inferences. In line with this, it has been shown before that asking inference questions about a text encourages readers to construct a situation model and learn new information (e.g. Roscoe & Chi, 2008). Under this scenario, any differences in text comprehension that might have shown up during reading (and were reflected in the associated visual processing measures) were washed out during the test phase when answering the comprehension questions.

Finally, the absence of an effect on reading comprehension scores may also be explained by the fact that the text and picture were in third-person perspective. Spatial perspective taking is crucial in understanding the viewpoint and perceptions of the protagonist (e.g. Rall & Harris, 2000). Moreover, text information is more readily available in the reader's mental representation when it is spatially close to the protagonist than when it is spatially separated from the protagonist (Morrow, Bower, & Greenspan, 1990). Both the present narrative text and the picture, however, were presented in a third-person perspective. This may have encouraged participants to adopt a similar perspective, rather than a first-person perspective (Salem, Weskott, & Holler, 2017). The participants in our study may have used the third-person perspective picture as a starting point for their situation model, resulting in more resources allocated to situation-model processes, but the discrepancy between their perspective and the protagonist's perspective could have impeded their reading comprehension performance. More research is needed to explore whether additional support (e.g. combining verbal instruction with one or multiple pictures) increases reading comprehension outcomes and to confirm the hypothesis that pictorial support that specifically encourages adopting a first-person perspective affects both process- and product-oriented measures of reading comprehension in a positive way.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Stephanie I. Wassenburg http://orcid.org/0000-0002-9007-3136
Björn B. de Koning http://orcid.org/0000-0001-5136-2261
Menno van der Schoot http://orcid.org/0000-0002-4762-2718

References

Aaronson, D., & Scarborough, H. S. (1977). Performance theories for sentence coding: Some quantitative models. *Journal of Verbal Learning and Verbal Behavior*, *16*, 277–303. doi:10.1016/S0022-5371(77)80052-2

Andreassen, R., & Bråten, I. (2011). Implementation and effects of explicit reading comprehension instruction in fifth-grade classrooms. *Learning and Instruction*, *21*, 520–537. doi:10.1016/j.learninstruc.2010.08.003

Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology*, *29*, 137–164. doi:10.1080/02702710801963951

Cain, K., & Oakhill, J. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, *11*, 489–503. doi:10.1023/A:1008084120205

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*, 5–26. doi:10.1023/A:1013176309260

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, *20*, 111–122. doi:10.1016/j.learninstruc.2009.02.010

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:10.1007/s10648-013-9222-6

Eitel, A., & Scheiter, K. (2015). Picture or text first? Explaining sequence effects when learning with pictures and text. *Educational Psychology Review*, *27*(1), 153–180. doi:10.1007/s10648-014-9264-4

Eitel, A., Scheiter, K., Schüler, A., Nyström, M., & Holmqvist, K. (2013). How a picture facilitates the process of learning from text: Evidence for scaffolding. *Learning and Instruction*, *28*, 48–63. doi:10.1016/j.learninstruc.2013.05.002

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading*, *17*(6), 435–452. doi:10.1080/10888438.2013.774005

Fletscher, C. R. (1994). Levels of representation in memory for discourse. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 589–608). San Diego: Academic Press.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, *31*, 129–151. doi:10.1016/0749-596X(92)90008-L

Graesser, A. C., Hoffman, N. L., & Clark, L. F. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior*, *19*, 135–151. doi:10.1016/S0022-5371(80)90132-2

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371–395. doi:10.1037/0033-295X.101.3.371

Guthrie, J. T., McRae, A., Coddington, C. S., Lutz Klauda, S., Wigfield, A., & Barbosa, P. (2009). Impacts of comprehensive reading instruction on diverse outcomes of low- and high-achieving readers. *Journal of Learning Disabilities*, *42*, 195–214. doi:10.1177/0022219408331039

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oup.

Houtveen, A. A. M., & van de Grift, W. J. C. M. (2007). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement*, *18*, 173–190. doi:10.1080/09243450601058717

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods*, *42*, 643–650.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209–226). Malden, MA: Blackwell.

Kraal, A., Koornneef, A. W., Saab, N., & van den Broek, P. W. (2018). Processing of expository and narrative texts by low- and high-comprehending children. *Reading and Writing*, *31*, 2017–2040. doi:10.1007/s11145-017-9789-2]

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge: University Press.

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 297–384). Burlington: Elsevier Inc. doi:10.1016/S0079-7421(09)51009-2

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

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1990). Situation-based inferences during narrative comprehension. *Psychology of Learning and Motivation*, *25*, 123–135. doi:10.1016/S0079-7421(08)60252-2

Newton, D. P. (1994). Pictorial support of discourse comprehension. *British Journal of Educational Psychology*, *64*, 221–229. doi:10.1111/j.2044-8279.1994.tb01097.x

Nyhout, A., & O'Neill, D. K. (2017). Children's enactment of characters' movements: A novel measure of spatial situation model representations and indicator of comprehension. *Mind, Brain, and Education*, *11*, 112–120. doi:10.1111/mbe.12142

Olsson, P. (2007). Real-time and offline filters for eye tracking (Unpublished master's thesis). Royal Institute of Technology, Stockholm, Sweden.
Pike, M. M., Barnes, M. A., & Barron, R. W. (2010). The role of illustrations in children's inferential comprehension. *Journal of Experimental Child Psychology*, *105*, 243–255. doi:10.1016/j.jecp.2009.10.006
Pouw, W., Rop, G., de Koning, B., & Paas, F. (2019). The cognitive basis for the split attention effect. *Journal of Experimental Psychology: General*. doi:10.1037/xge0000578
Radvansky, G. A., Zwaan, R. A., Curiel, J. M., & Copeland, D. E. (2001). Situation models and aging. *Psychology and Aging*, *16*(1), 145–160. doi:10.1037//0882-7974.16.1.145
Rall, J., & Harris, P. L. (2000). In Cinderella's slippers? Story comprehension from the protagonist's point of view. *Developmental Psychology*, *36*, 202–208. doi:10.1037/0012-1649.36.2.202
Rayner, K. (1998). Eye movements in reading and information processing: 20 Years of research. *Psychological Bulletin*, *124*, 372–422. doi:10.1037/0033-2909.124.3.372
Rinck, M., & Bower, G. H. (1995). Anaphora resolution and the focus of attention in situation models. *Journal of Memory and Language*, *34*(1), 110–131. doi:10.1006/jmla.1995.1006
Roscoe, R. D., & Chi, M. T. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science*, *36*(4), 321–350. doi:10.1007/s11251-007-9034-5
Salem, S., Weskott, T., & Holler, A. (2017). Does narrative perspective influence readers' perspective-taking? An empirical study on free indirect discourse, psycho-narration and first-person narration. *Glossa: A Journal of General Linguistics*, *2*(1), 61. doi:10.5334/gjgl.225
Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, *14*(1), 101–120. doi:10.1023/A:1013136727916
Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*, 141–156. doi:10.1016/S0959-4752(02)00017-8
Schüler, A., Arndt, J., & Scheiter, K. (2015). Processing multimedia material: Does integration of text and pictures result in a single or two interconnected mental representations? *Learning and Instruction*, *35*, 62–72. doi:10.1016/j.learninstruc.2014.09.005
Stine, E. L. (1990). On-line processing of written text by younger and older adults. *Psychology and Aging*, *5*(1), 68–78. doi:10.1037//0882-7974.5.1.68
Stine-Morrow, E. A. L., Gagne, D. D., Morrow, D. G., & DeWall, B. H. (2004). Age differences in rereading. *Memory and Cognition*, *32*, 696–710. doi:10.3758/BF03195860
Stine-Morrow, E. A. L., Milinder, L.-A., Pullara, O., & Herman, B. (2001). Patterns of resource allocation are reliable among younger and older readers. *Psychology and Aging*, *16*(1), 69–84. doi:10.1037/0882-7974.16.1.69
van den Broek, P. (1997). Discovering the cement of the universe: The development of event comprehension from childhood to adulthood. In P. van den Broek, P. J. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension and representation: Bridging fictional and actual events* (pp. 321–343). Mahwah, NJ: Erlbaum.
van den Broek, P., & Lorch, R. F. (1993). Network representations of causal relations in memory for narrative texts: Evidence from primed recognition. *Discourse Processes*, *16*, 75–98. doi:10.1080/01638539309544830
van der Schoot, M., Horsley, T. M., & van Lieshout, E. C. D. M. (2010). The effects of instruction on situation model construction: An eye fixation study on text comprehension in primary school children. *Educational Psychology*, *30*, 817–835. doi:10.1080/01443410.2010.510600
Wassenburg, S. I., Beker, K., van den Broek, P., & van der Schoot, M. (2015). Children's comprehension monitoring of multiple situational dimensions of a narrative. *Reading and Writing*, *28*, 1203–1232. doi:10.1007/s11145-015-9568-x
Wassenburg, S. I., Bos, L. T., de Koning, B. B., & van der Schoot, M. (2015). Effects of an inconsistency-detection training aimed at improving comprehension monitoring in primary school children. *Discourse Processes*, *52*, 463–488. doi:10.1080/0163853X.2015.1025203
Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *The Journal of General Psychology*, *132*, 408–428. doi:10.3200/GENP.132.4.408-428

Wolfe, M. B. W. (2005). Memory for narrative and expository text: Independent influences of semantic associations and text organization. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*, 359–364. doi:10.1037/0278-7393.31.2.359
Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1196–1207. doi:10.1037/0278-7393.22.5.1196
Zwaan, R. A. (1999). Situation models. *Current Directions in Psychological Science*, *8*(1), 15–18. doi:10.1111/1467-8721.00004
Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*, 292–297. doi:10.1111/j.1467-9280.1995.tb00513.x
Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 386–397. doi:10.1037//0278-7393.21.2.386
Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185. doi:10.1037/0033-2909.123.2.162

Appendix

Original version of the story (English translation below):

De reis van de muis

In de stad Weesp woonde een muis die op een maandagavond in de trein sprong.
Hij wilde zijn nicht bezoeken.
De muis reed mee tot Baarn en kroop in de jas van een man, die in de bus stapte.
De bus ging naar De Bilt en daar sprong de muis op een auto.
Maar die reed zo hard dat hij eraf woei door de wind, juist op een brug waar een boot onderdoor voer.
De muis viel in de boot en kwam de volgende ochtend aan in Tiel.
Daar woonde zijn nicht, in het huis van de dokter.
Nadat hij aangeklopt had bij haar holletje deed zij de deur open.
Zijn nicht was verbaasd hem te zien en vroeg hoe hij hier gekomen was.

De muis vertelde dat hij met de trein en de bus en de auto en de boot had gereisd.
Hij bleef twee dagen logeren bij zijn nicht.
Wanneer het dag was, aten ze van de kaas in de keuken van de dokter.
En in de nacht dansten ze op de tafel die gedekt was.
Hup-twee hup-twee langs de pot met jam en de pot met stroop en de pot met honing.
Daarna sliepen ze in het servet van de dokter als in een bed met witte lakens.
Donderdagochtend moest de muis weer naar huis.
Zijn nicht was benieuwd hoe hij nu weer in Weesp zou komen.
De muis dacht even na en zei toen dat hij het zou proberen met de post.
Hij gaf haar een kus en sprong in een doos, die klaar lag om te worden gepost.

De doos ging naar Den Haag, maar in de trein kroop de muis eruit en zocht een pak voor Weesp.
Hij sprong in het vak met een W erop.
Daar lag een pak voor Weert, voor Wijk en voor Wilp, maar geen voor Weesp.
Dan maar naar Baarn, dacht de muis.
Maar in het vak met de B lag alleen maar een pak voor Bussum.
Dat is ook goed, dacht de muis, want Bussum ligt in de buurt van Weesp.
Hij beet een gat in het pak en kroop erin.
In Bussum sprong hij uit de trein en liep door naar de sloot, daar lag een klomp.

Hij maakte een zeil van een krant en voer in een middag naar Weesp.
Thuis zette hij een nieuw naambordje op de deur van zijn hol: J. Muis, ontdekkingsreiziger.

English translation:
The journey of the mouse
In the city of Weesp there lived a mouse who jumped on a train on Monday evening.
He wanted to visit his cousin.
The mouse rode along until Baarn and crawled inside a man's jacket, who then got on a bus.
The bus went to De Bilt, and there, the mouse jumped onto a car.
But it went so fast that he was blown off by the wind onto a bridge, while a boat was sailing underneath.

The mouse fell into the boat and arrived in Tiel the next morning.
His cousin lived there, in the doctor's house.
After he knocked at her mousehole, she opened the door.
His cousin was surprised to see him and asked him how he had got here.
The mouse said that he had traveled on the train, the bus, the car, and the boat.
He stayed at his cousin's for two days.
During the day, they ate cheese in the doctor's kitchen.
And during the night, they danced on the table that was set.
Hop-two hop-two, along the jam jar, and the syrup jar, and the honey jar.
Afterwards, they slept using the doctor's napkin as a bed with white sheets.
On Thursday morning, the mouse had to go back home.
His cousin was curious about how he would get back to Weesp.
The mouse thought for a moment and said that he would try to do it by mail.
He gave her a kiss and jumped into a box that was waiting to be posted.

The box was going to Den Haag, so on the train, the mouse crawled out and looked for a package to Weesp.
He jumped into the box that had a W on it.
There was a package for Weert, for Wijk, and for Wilp, but none for Weesp.
Then let's go to Baarn, the mouse thought.
But in the box with a B there was only a package for Bussum.
That's also fine, the mouse thought, because Bussum is close to Weesp.
He bit a hole in the package and crawled inside.
In Bussum, he jumped off the train and walked to a ditch where there was a clog.
He made a sail from a newspaper and sailed to Weesp in an afternoon.
Back home, he put a nameplate on the door of his mousehole: J. Mouse, the explorer.