Cognitive processes underlying the text-belief consistency effect: An eye-movement study

Johanna Maier & Tobias Richter
University of Würzburg

M. Anne Britt
Northern Illinois University

Manuscript accepted for publication in the journal Applied Cognitive Psychology (2018)

Author Note

The research reported in this article was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG, grant RI 1100/5-2 and grant RI 1100/5-3). We would like to thank Veronika Schmitt, Caroline Ahrendts, and Sandra Opoka for their help in collecting and cleaning the eye-tracking data. The experimental texts, data files, and R-scripts for the full analyses are available from the authors upon request.

Corresponding author:
Johanna Maier
University of Würzburg, Department of Psychology IV
Röntgenring 10
97070 Würzburg, Germany
E-mail: johanna.maier@uni-wuerzburg.de

Abstract

Readers' memory for belief-consistent texts is often stronger than for belief-inconsistent texts (text-belief consistency effect). However, presenting belief-consistent and belief-inconsistent texts alternatingly reduces the discrepancy between the memory strengths of belief-consistent and belief-inconsistent texts. The present study used eye tracking to examine the cognitive processes underlying the text-belief consistency effect and how it is moderated by the mode of presentation. At two points of measurement, forty-one university students read two belief-consistent and two belief-inconsistent texts on two scientific issues blocked or alternatingly. Comprehension outcomes were assessed with an essay task. First-pass rereading times were longer for belief-inconsistent information for participants with strong beliefs. A blocked presentation increased this effect and yielded longer first-pass rereading times for belief-inconsistent claims and a text-belief consistency effect in the essay task. An alternating presentation increased immediate and delayed processing of belief-inconsistent information and reduced the text-belief consistency effect, especially in readers making more lookbacks.

Keywords: multiple texts, beliefs, mode of presentation, eye-tracking.

Cognitive processes underlying the text-belief consistency effect: An eye-movement study

The World Wide Web has become one of the primary sources of science-related information. It offers great possibilities such as the easy availability of new scientific results, but it also provides a challenge to readers who are faced with a multitude of complex and often inconsistent texts. Therefore, readers need to be able to understand and critically evaluate scientific content to achieve their goals (Britt, Richter, & Rouet, 2014). However, many readers experience difficulties in judging the acceptability of scientific claims from a neutral point of view, which may impact their comprehension of scientific discourse. Hotly debated science-related topics, such as the causes of global warming, are characterized by strongly held beliefs of readers on either side of the scientific controversy. Research suggests that readers' prior beliefs play an important role in the learning and retention of information (e.g., Kardash & Scholes, 1996; Knobloch-Westerwick & Meng, 2011; van Strien, Brand-Gruwel, & Boshuizen, 2014). Many studies have shown that individuals maintain their beliefs even in the face of new information that explicitly corrects or discredits them (e.g., Chinn & Brewer, 1993; Limon & Mason, 2002; Johnson & Seifert, 1994; Ross, Lepper, & Hubbard, 1975). Likewise, readers' comprehension of controversial topics is often biased by prior beliefs leading to a stronger representation of belief-consistent compared to belief-inconsistent information (*text-belief consistency effect*; e.g., Eagly & Chaiken, 1993; Maier & Richter, 2013, 2014; Wiley, 2005). Maier and Richter (2013) examined the impact of prior beliefs on the memory representation of multiple conflicting texts. In their study, participants read two belief-consistent and two belief-inconsistent texts either in a blocked or an alternating mode of presentation. After participants had read one text, their comprehension of the state of affairs described in the text (*situation model*; Kintsch, 1998) was assessed with a recognition task with inference items. Readers' situation model was stronger for belief-consistent texts compared to the situation model of the belief-inconsistent texts but only when participants read the texts in a blocked mode of presentation.

However, despite the fact that text-belief consistency effects are well established, little is known about the cognitive processes underlying such effects. The present research advances this line of research by using eye tracking as a means of examining moment-to-moment processes during reading belief-consistent and belief-inconsistent texts. This method allows a non-intrusive and immediate investigation of information processing (Just & Carpenter, 1980). Moreover, text-based learning is not only influenced by reader characteristics but is a rather complex interplay of the reader and the reading situation, for instance the way texts are structured and presented to readers. In our study, we were interested in how the sequence of presenting multiple contradictory texts about a scientific controversy influences processing and comprehension of belief-consistent and belief-inconsistent texts. For this aim, the texts were either presented in an alternating (i.e., belief-consistent text followed by belief-inconsistent text) or in a blocked mode of presentation (i.e., belief-consistent text followed by another belief-consistent text). After reading, readers' memory of the texts was assessed with an essay task.

In the following sections, we first discuss the role of prior beliefs in the processing of arguments, which may be regarded as the most important building blocks of scientific texts. Against this background, we will discuss potential differences in the processing of belief-consistent and belief-inconsistent texts and argue why the sequence in which such texts are encountered might moderate the text-belief consistency effect.

Arguments in Scientific Texts

Texts on disputed science topics that are available on the Internet often present one-sided arguments intended to support one particular argumentative position. In line with Kintsch (1998), the argumentative position of a text can be viewed as part of the texts' macrostructure. In detail, the macrostructure can be viewed as a semantic structure that depicts at a more global level a texts' overall unity and coherence and includes aspects such as topic or theme. For instance, a texts' title or subheadings within a text provides macrostructure cues that state a text's position. In the case of one-

sided arguments, readers should be able to conclude the texts' argumentative position based on such macrostructural cues.

If a reader holds pertinent beliefs, the arguments presented in the text can be consistent or inconsistent with such beliefs. Toulmin (1958) proposed that a minimal argument consists of a claim that is supported by data (i.e., reasons, Voss, Fincher-Kiefer, Wiley, & Silfies, 1993). Based on Toulmin's model, Voss et al. (1993) examined how the processing of claim-first arguments (i.e., arguments in which the claim is presented first followed by supporting reason) is influenced by beliefs. In particular, Voss and colleagues argued that presenting a claim is sufficient to activate beliefs, which are then used to assess the truth of the claim. This relationship depends on belief-topic strength with faster judgments occurring when readers have stronger beliefs. In line with these assumptions, Voss et al. found that for participants with strong beliefs, agree-disagree judgments for sentences were as fast as judgments of meaningfulness. In a second experiment, participants read a combination of claim and reason. The reason was varied in four conditions as strong vs. weak and supporting vs. opposing. When judging whether or not the reason was supportive or opposed to the claim, participants were faster in responding to reasons that were in line with their beliefs and that were strong and supporting reasons. These results suggest that a claim activates prior beliefs quickly during reading. Furthermore, claim-belief consistency seems to play a critical part in how claims are processed. Additional support for this conclusion comes from a study on argument evaluation by Wolfe, Britt and Butler (2009, Study 2). In their study, participants judged agreement and argument quality after reading one-sided arguments that contained either a belief-consistent or a belief-inconsistent claim and either a similar or a dissimilar reason. The authors found evidence that agreement with an argument was mainly driven by agreement with the claim (for similar results, see von der Mühlen, Richter, Schmid, Schmidt, & Berthold, 2016). Moreover, whereas claim agreement influenced mainly argument agreement, Wolfe

and colleagues found (2009, Study 2) that characteristics of reasons had a stronger influence on perceived argument strength or quality.

Selective Exposure versus Disconfirmation

How belief-consistent vs. belief-inconsistent information is processed might also depend on different attention foci. On the one hand, superior memory for belief-consistent information found in text-belief consistency studies might be a result of paying more attention to belief-consistent information (a type of selective exposure effect, Festinger, 1957, or congeniality bias, Eagly & Chaiken, 1993). In line with this assumption, research on information selection has found that readers have a tendency to select information that supports their viewpoint and to avoid information that might challenge their beliefs (Hart, Albarracin, Eagly, Brechan, Lindberg, & Merrill, 2009; Jonas, Schulz-Hardt, Frey, & Thelen, 2001; Knobloch-Westerwick & Meng, 2011). Such a preference for belief-consistent information can be viewed as a defense motivation to avoid cognitive dissonance (Festinger, 1957). A similar bias might occur in the allocation of attention to belief-consistent and belief-inconsistent information during reading – especially when avoiding belief-inconsistent information is not possible.

In contrast to this view, the disconfirmation model proposed by Edwards and Smith (1996) assumes that readers pay more attention to belief-inconsistent information. Edwards and Smith suggested that belief-consistent arguments are accepted at face value, whereas belief-inconsistent arguments set a deliberative memory search in motion with the goal to undermine the arguments and to explain why they are incorrect. This type of processing can be viewed as an active defense against attack mechanism, causing readers to allocate an extra amount of processing resources on belief-inconsistent information to thoughtfully counterargue it (Eagly, Kulsesa, Brannon, Shaw, & Hutson-Comeaux, 2000).

In sum, readers can either protect their beliefs by ignoring or by devaluating belief-inconsistent information (both are types of *belief protection*, McCrudden & Sparks, 2014) or readers can use belief-inconsistent information to update or modify their prior beliefs (*belief reflection;* McCrudden & Sparks, 2014). Empirical research is still needed to determine the circumstances under which belief-consistent or belief-inconsistent information receives more attention during immediate and delayed processing.

Prior Beliefs and Processing

In research on relevance (e.g., McCrudden, Magliano, & Schraw, 2011) and perspective effects (e.g., Anderson & Pichert, 1978) it is assumed that a reading perspective defines which information in a text is relevant and which is irrelevant. Moreover, information perceived as relevant for a given perspective or reading goal receives more attention during reading and is better remembered after reading (Anmarkrud, McCrudden, Bråten, & Strømsø, 2013; Burton & Daneman, 2007; Kaakinen & Hyönä, 2005, 2008, 2011; Kaakinen, Hyönä, & Keenan, 2003; McCrudden, Magliano, & Schraw, 2010). For instance, Kaakinen and colleagues (2003) presented participants an expository text that contained an equal amount of sentences providing information about two illnesses. Prior to reading, participants were instructed to find out facts about one of the two diseases. This instruction was used to induce a reading perspective. An investigation of participants' eye movements revealed longer processing of perspective-relevant information for indicators of immediate and delayed processing. In addition, a similar perspective effect was observed in recall data with higher recall of perspective-relevant information.

It seems possible that readers' prior beliefs can also serve as standards of relevance and thus, might determine the instrumental value of a text segment in relation to readers' reading task. This would suggest that prior beliefs are also able to affect immediate and delayed processing. However, one important difference between relevance effects and the influence of prior beliefs on reading is that

relevance instructions are directly given prior to reading, whereas readers might not be fully in active control of the influence of their prior beliefs. Rather, the activation of prior beliefs during comprehension is partly due to memory-based processes in which relevant information is retrieved quickly and unrestricted from long-term memory (e.g., O'Brien & Myers, 1999; O'Brien & Cook, 2015). Hence, all information that resonates sufficiently during reading is returned to working memory, regardless of its relevance. Strong prior beliefs are likely to be also activated and retrieved in this resonance process, regardless of whether they hinder or facilitate comprehension. In addition, once activated and retrieved, prior knowledge and beliefs can be used to validate text information (Richter, 2015; Singer, 2006), which can occur immediately or delayed during reading. If validation occurs immediately during comprehension, the processing includes an involuntary check of the consistency of text information with prior knowledge and beliefs (e.g., Richter, Schroeder, & Wöhrmann, 2009; Singer, 2006; for a detailed discussion see Isberner & Richter, 2014). Information that fails the consistency check is less likely to be integrated into the situation model of the text content (Schroeder, Richter, & Hoever, 2008). Consequently, when prior beliefs are used to validate information immediately during processing, this might result in processing difficulties for belief-inconsistent claims (Wolfe et al., 2009) and weaker memory for belief-inconsistent information.

Moreover, if immediate validation fails to result in a mental representation that readers perceive as sufficiently coherent, additional processes directed at establishing coherence may be initiated (van den Broek, Beker, & Oudega, 2015). As a result, readers can also engage in a more delayed validation process that includes strategic memory retrieval and a complete evaluation of all information (Maier & Richter, 2016; Richter, 2015). Such a type of validation requires delayed knowledge-based processing directed at resolving contradictions between texts (such as lookbacks to previous read text parts).

Mode of Presentation

In informal learning, readers can encounter multiple belief-consistent and belief-inconsistent texts in different sequences. They can read texts of the same belief type successively, or they can encounter texts of different belief types alternately. For example, individuals might first read a belief-consistent text followed by a belief-inconsistent text before they again read a belief-consistent text. Such an alternating mode of presenting belief-consistent and belief-inconsistent information—in contrast to a blocked mode of presentation—has been shown to reduce the memory advantage for belief-consistent information (Maier & Richter, 2013; Wiley, 2005).

To our knowledge, no research to date has directly investigated differences in cognitive processing of belief-consistent and belief-inconsistent multiple texts when the texts are presented alternately vs. blocked. Research supports, however, the assumption that the influence of prior beliefs on the processing of conflicting information depends on text properties (Braasch, Goldman, & Wiley, 2013; Kendeou, Walsh, Smith & O'Brien, 2014; Maier & Richter, 2013; Wiley, 2005). In particular, research in educational psychology suggests that the impact of readers (false) prior beliefs can be reduced by refutational texts (Ariasi & Mason, 2011; Braasch et al., 2013; Kendeou et al., 2014). Kendeou and O'Brien (2014) proposed that refutational texts are especially beneficial in revising false beliefs, because "newly acquired information makes contact with, and activates, the preexisting knowledge base so that both pieces of information are in working memory at the same time" (*coactivation*; Kendeou & O'Brien, 2014, p. 353). When this coactivation occurs, newly encoded information can be integrated with previous knowledge and beliefs. A similar mechanism might underlie the capability of an alternating mode of presentation to reduce the text-belief consistency effect. In an alternating mode of presentation, concepts stated in the two different text types are more likely to be co-activated, resulting in increased awareness of intertextual inconsistencies and an increased need for coherence-based processes (van den Broek et al., 2015) to resolve or explain the inconsistency. Put differently, an alternating mode of presentation might help readers in detecting

intertextual inconsistencies, which should lead to an increase in the need for repair processes to resolve the inconsistency.

Rationale and Overview of the Present Experiment

Previous research indicates that a reader's prior beliefs can bias the comprehension of multiple controversial texts to the extent that belief-consistent texts are comprehended better compared to belief-inconsistent texts (e.g., Maier & Richter, 2013). However, presenting belief-consistent and belief-inconsistent texts (or arguments) in an alternating manner seems to reduce the text-belief consistency effect (Maier & Richter, 2013; Wiley, 2005). However, the processes underlying the text-belief consistency effect and the moderating role of the presentation mode are unclear. The aim of the present study was to advance the literature by investigating the on-line processing of belief-consistent and belief-inconsistent multiple texts that were read alternately or block-wise. Learners' eye-movement patterns for sentences as units of analysis during reading belief-consistent and belief-inconsistent texts either in an alternating or in a blocked mode of presentation were recorded at two times of measurement.

We were interested in the extent that text-belief consistency and mode of text presentation influence first- and second-pass fixation measures (Hyönä, Lorch, & Rinck, 2003) for claims and reasons, which are the central functional elements of an argument (Toulmin, 1958). First-pass reading time and first-pass rereading time are two commonly reported indicators of immediate processing when sentences are the unit of analysis (Hyönä et al., 2003). *First-pass reading time* is the summed duration of the first-pass fixations during initial reading of a sentence before moving on or moving back to earlier parts of the sentence or text. Put differently, first-pass reading time includes all fixations within a sentence as long as the eye moves forward along the sentence. *First-pass rereading time* is based on the "duration of all reinspective fixations landing on the target region during its first-pass reading" (Hyönä et al., 2003, p. 332). It includes the fixations that occur if the eyes return to already

read parts within a sentence (before this sentence is exited) and can be viewed as indicating readers' need to immediately reread a sentence (Kaakinen et al., 2003). Hence, in contrast to first-pass reading time, which is often not sensitive to inconsistency effects (Rinck, Gamez, Diaz, & de Vega, 2003), first-pass rereading time reliably indicates comprehension difficulties during initial reading. In other words, disruption in processing during initial reading becomes visible in longer first-pass rereading times. The second-pass fixation measure investigated in our study was the number of lookbacks, which indicates delayed processing (Hyönä et al., 2003). Lookbacks index whether the reader returns to sentences from a subsequent sentence, which indicates that readers reinstate information to the working memory (Walczyk & Taylor, 1996) and look for information that might resolve or explain an inconsistency (Rinck et al., 2003). However, such repair processes are thought to only occur if inconsistencies are consciously noted (Rinck et al., 2003).

Assumptions for first-pass reading times

As first-pass reading time is often not sensitive to inconsistency effects (Rinck, Gamez, Diaz, & de Vega, 2003), we did not expect any effects of text-belief consistency on first-pass reading times.

Assumptions for first-pass rereading times

We expected that when readers monitor belief-consistency immediately during reading, longer first-pass rereading time for belief-inconsistent information should occur as such information should increase the need to immediately reread a sentence. However, considering that Voss et al. (1993) found that the strength of readers' prior beliefs is a crucial determinant for the monitoring of belief-consistency, we assumed that longer first-pass rereading times for belief-inconsistent information should occur only for readers with strong prior beliefs (Hypothesis 1). Moreover, we expected this immediate monitoring effect of belief-consistency for strong believers to be moderated by mode of presentation. An alternating mode of presentation – similar to refutational texts – might increase readers' awareness of intertextual inconsistencies, whereas a blocked mode of presentation should

make this more difficult. Hence, we expected longer first-pass rereading times for belief-inconsistent information in high believers (i.e., participants that have strong prior beliefs) in a blocked mode of presentation (Hypothesis 2).

Moreover, we were expecting that the immediate need to reread belief-consistent and belief-inconsistent claims and reasons should differ as a function of the presentation mode. If the assumption holds that belief-consistency is relevant especially during claim processing (Voss et al., 1993; Wolfe et al., 2009), first-pass rereading time for belief-inconsistent claims should be longer when controversial texts are presented in a blocked mode compared to an alternating mode of presentation (Hypothesis 3). The predictions for eye fixations on reasons are not so straightforward. As opposed to claims, reasons are not directly consistent or inconsistent with readers' prior beliefs. Thus, an immediate monitoring of text-belief consistency should not affect the processing of reasons. Understanding reasons, however, is crucial for understanding the argument. Based on this consideration, we assumed that in a blocked mode of presentation, in which belief-consistent information is expected to be perceived as more relevant, first-pass rereading time for belief-consistent reasons should be more pronounced (Hypothesis 4a). In an alternating mode of presentation, this pattern is expected to be reversed—first-pass rereading time for belief-inconsistent reasons is expected to be more pronounced, because understanding the reasons for belief-inconsistent claims is crucial for defending beliefs (Hypothesis 4b).

Assumptions for lookbacks

In contrast to a blocked mode of presentation, an alternating mode of presentation was expected to enhance repair processes and search for additional information especially when reading belief-inconsistent information, because inconsistencies between texts should be more salient in this mode of presentation. As an indicator of such repair processes, we expected readers to show more *lookbacks* for

belief-inconsistent information when reading belief-consistent and belief-inconsistent texts alternately (Hypothesis 5).

Assumptions for essay task

In addition to examining underlying cognitive processes, we purposed to replicate the text-belief consistency effect in the comprehension outcome found by Maier and Richter (2013, 2014) with a different task. The previous studies used a recognition or verification task (adapted from Schmalhofer & Glavanov, 1986) to measure readers' comprehension of each text. This relatively simple method for estimating mental model strength might not capture more advanced aspects of mental model quality. The present study used an essay task in which participants were asked to write an argumentation after all texts were read (Anmarkrud et al., 2013). The advantage of this task over the recognition/verification task is that participants need to generate and relate arguments presented in the texts to justify their own position. The task requires a deep understanding of the text and combining text information (from both sides of the issue) with prior knowledge. In that sense, it may be regarded as a way to measure text comprehension on the level of the mental model of the controversy. However, given that participants are asked to formulate and justify their position on the controversy, it also involves readers' (post-reading) beliefs to a stronger degree than the tasks used in previous research. For essay tasks of this kind, text-belief consistency effects have been found in previous research on multiple text comprehension (see Richter & Maier, 2017, for an overview), which may be regarded as an instance of the myside bias (Toplak & Stanovich, 2003; Wolfe et al., 2009). Thus, for the essay task, we expected participants to generate more belief-consistent arguments overall as a result of the myside bias (Hypothesis 6). However, this text-belief consistency effect was expected to be reduced by an alternating compared to a blocked mode of presentation (Hypothesis 7).

Finally, we were also interested in whether or not participants' individual effects in delayed processing would moderate the assumed relationship between text-belief consistency and mode of text

presentation for the learning outcome. Based on the assumption that an alternating text presentation will reduce the text-belief consistency effect by encouraging delayed processing of belief-inconsistent information, participants with a higher amount of lookbacks (as indicator of delayed processing) should especially benefit from an alternating mode of text presentation (Hypothesis 8).

Method

Participants and Manipulation Check for Text-Belief Consistency

Forty-two university students took part at the study and were first screened for prior beliefs (see section "prior beliefs" for details). Overall, participants more strongly agreed with the stance of the pro texts for both topics. Paired sample *t* tests revealed that for global warming, agreement to the argumentative position that mankind caused global warming was stronger ($M = 4.62$, $SD = 1.03$) compared to the agreement with the position that global warming is a result of natural phenomena ($M = 1.98$, $SD = 0.88$), $t(42) = 10.04$, $p < .001$, $d = 2.77$. Similarly, participants more strongly agreed with the argumentative position that vaccinations are necessary and beneficial ($M = 4.32$, $SD = 0.91$) than with the position that vaccinations are unnecessary and harmful ($M = 2.19$, $SD = 0.81$, $t(42) = 9.21$, $p < .001$, $d = 2.48$. However, a detailed investigation of responses revealed that five participants more strongly agreed to the argumentative stance of the vaccination contra texts, that is, vaccinations are unnecessary and harmful. Similarly, four participants more strongly agreed to the contra stance that natural phenomena caused global warming. One participant preferred for both topics the contra argumentative stance. These uneven cell sizes for both topics made it impossible to balance out belief-consistency. For this reason, only data points where participants favored the stance of the pro texts for the two topics, that is, mankind caused global warming and vaccinations are beneficial, were used in

further analyses and this argumentative position is further referred to as belief-consistent position.[1] The argumentative stance of the contra text, in turn, is referred to as belief-inconsistent position.

The remaining sample consisted of forty-one psychology students (33 women and 8 men) with an average age of 24.10 (*SD* = 7.57) and an average semester of 2.46 (*SD* = 1.48). Students received course credit for participating.

Apparatus

Participants' eye movements were recorded with a desk-mount EyeLink 1000 eye tracker manufactured by SR Research Ltd. (Ontario, Canada) with a chin-forehead rest. Pupil location was sampled at a rate of 1000 Hz with an average spatial accuracy of 0.5º. Movements of participants' dominant eye (usually the right eye) were recorded, but viewing was binocular. The text material was presented on a 22'' TFT monitor with a resolution of 1024 x 768 and a refresh rate of 75 Hz. Viewing distance was approximately 86 cm.

Text Material

The text material consisted of eight generally accessible texts for two scientific controversial topics (global warming/vaccinations) and was based on text material used in a study from Maier and Richter (2013). For the scientific controversy about the risks and benefits of vaccinations, two texts argued in favor of vaccinations (pro-text stance) and two texts argued against vaccinations (contra-text stance). Similarly, for the texts on global warming, two texts took the stance that global warming is caused by mankind (pro-text stance) and two texts took the stance that global warming is caused by natural phenomena (contra-text stance). The writing style, structure, and length of the text were held strictly parallel (see Table 1 for a synopsis of text characteristics). The average length of the text was

[1] This data cleaning strategy led to the fact that for 25 participants data from both text topics (in different orders and modes of presentation) could be analyzed, whereas for 16 participants this was true for only one eye-tracking session. This approach seems appropriate as linear mixed models allow including incomplete data sets and still ensured a sufficient number of data points.

606 words. The average readability score (determined with the German adaption of the Flesch's Reading Ease Index, Amstad, 1978) was 44, indicating a moderate difficulty. All texts consisted of six parts (see Figure 1 for an example in original formatting). The texts started with a short introduction in which the scientific topic and the argumentative stance of the text were introduced. Each text then presented four unique arguments. Each argument had a sub headline and presented the claim in the first sentence of the argument. Claims were unique for each argument and were further matched for mean word frequency and lengths. In the remaining sentences of the argument, the reason was presented. Each reason comprised of 4 to 9 sentences (M = 6.7, SD = 1.4). The total number of claims (4) and reasons (26-27) was held constant between texts. The texts ended with a short summary and a conclusion. To ensure that no differences in text quality were perceived, the texts were pilot-tested with an independent sample of 45 university students. In this pilot test, participants read two texts and judged the text with regard to understandability, plausibility, interest, and number of arguments (see Table 1). Paired-samples t tests (with Holm-Bonferroni correction for multiple tests, Holm, 1979) revealed no significant differences in judgments between the texts. Thus, all eight texts can be viewed as equally understandable and interesting and to contain four arguments as expected.

Essay Task

In the essay task, participants received a summary of the scientific debate in form of a controversial question and were instructed to express their view on the topic. In detail, the instruction participants received in the essay task was the following (translated into English):

> After reading the texts, we ask you to make a statement about the scientific discourse that was presented in the texts. Please answer the following question: "Is global warming a natural phenomenon or is mankind responsible for global warming?/ Are vaccinations a good and important preventative measure or a threat to health?"[Depending on the text topic only one of these questions was presented]. Please justify your statement with as many arguments as

possible. It is important to cite arguments that support your position as well as arguments that run against your position.

Reader Characteristics

Prior beliefs. Participants' prior beliefs about the cause of global warming were assessed by nine statements and their prior beliefs about the risks and benefits of vaccinations were assessed by 10 statements (response categories ranging from 1 = *totally disagree* to 6 = *totally agree*). Five statements per scientific controversy assessed participants' agreement to the argumentative stance of the pro texts, that is, mankind is responsible for global warming (e.g., "I believe that humans are the cause of global warming", Cronbach's $\alpha = .83$) and vaccinations are more beneficial than harmful (e.g., "I think that vaccinations are the most important and most effective method against infectious diseases", Cronbach's $\alpha = .76$). Similarly, participants' agreement to the argumentative stance of the contra global warming texts (i.e., natural phenomena are the causes of global warming) were assessed with four items (e.g., "I believe that the climate on earth has always changed from time to time as long as the earth exists", Cronbach's $\alpha = .79$).[2] Participants' agreement to the argumentative stance of the contra vaccinations texts (i.e., that vaccinations are unnecessary and harmful) were assessed with five items (e.g., "I am against vaccinations, because they might overstrain my immune system", Cronbach's $\alpha = .77$).

Reading skills. We assessed participants' reading skills with the sentence verification subtest of the German reading test ELVES (Richter & van Holt, 2005, Cronbach's $\alpha = .87$) for control purposes. The ELVES assesses readers' efficiency in using propositional strategies for reading comprehension at the sentence level (i.e., lexical access; syntactic and semantic integration). Participants are required to

[2] Similar to the other belief scales, five statemens were given to participants to assess their agreement to the argumentative stance of the contra global warming texts. However, one item of this scale was excluded from the analyses because it was not or even negatively correlated with the other items from the scale.

judge the correctness of assertions about abstract and concrete concepts (true or false). Test scores combine the accuracy and the speed of a given response for each assertion (for details, see Richter & van Holt, 2005).

Procedure

Two weeks prior to the next session, reader' characteristics were measured to minimize carry-over effects. Eye movements were recorded in two sessions with a time-lag of two weeks. In each session, participants read the four texts on one scientific controversy in one mode of presentation. For example, if participants received belief-consistent and belief-inconsistent texts about one of the two scientific topics in an alternating mode of presentation in the first session, they received belief-consistent and belief-inconsistent texts of the other scientific topic block-wise in the second session. The order of the mode of presentation and the text topic were counterbalanced across participants. At the beginning of each reading session, participants were simply informed that they would read four texts on vaccination respectively global warming. Before participants read a text, a 9-point calibration grid covering the entire screen was used to calibrate the eye-tracker. The calibration procedure was repeated as often as necessary to achieve an average angular error of less than 0.5°. When the measurement was no longer accurate during reading, the calibration procedure was repeated. Participants read the texts at their own pace. On one screen, one segment of a text (introduction, argument 1-4, and conclusion) was presented. The texts were presented on the screen in black Courier New (16 point font for headings and 13.5 point font for main text, double-spaced) on a light grey background. A maximum of sixteen lines of text (including subtitle) were presented on one screen. Participants moved to the next text segment by pressing the space bar. After participants finished reading one text, they worked on comprehension questions (24 items per text; e.g., "Does this information match the situation described by the texts: Infamous and dangerous germs are defeated by

vaccinations.") to ensure that participants had understood the texts.[3] After all four texts were read, participants worked on the essay task and completed a reading strategy questionnaire (21 items; e.g., "I compared the arguments from the text with my opinion about the issue"). At the end of the experiment, participants were thanked and debriefed.

Design

The experimental design was a 2 (*text-belief consistency*: consistent vs. inconsistent) x 2 (*order of presentation*: block-by-block vs. alternating) x 2 (*argument part*: claim vs. reason) within-subjects design. In addition, the order of belief-consistent vs. belief-inconsistent texts (belief-consistent first vs. belief-inconsistent first), text topic (global warming vs. climate change), eye-tracking measurement time (session 1 vs. session 2), the texts itself, and participants' belief strength and reading skills were included as control factors or continuous covariates in the analysis.

Results

Descriptive Statistics for Belief Strength

For belief-strength, we computed a difference score as indicator (agreement to belief-consistent position – agreement to belief-inconsistent position). Overall, participants had a clear preference for the belief-consistent stance, but belief strength varied greatly (global warming: $M = 3.14$, $SD = 1.17$, range: 0.75-4.80; vaccination: $M = 2.57$, $SD = 0.96$, range: 0.40-4.60).

Effects of Text-Belief Consistency on Eye-Movement Measures

For the eye-movement measures, sentences were the level of analyses. Table 2 provides an overview of the number of sentences, that is trials, within the 2x2x2 Design (Topic x Text Stance x Text) for the eye-movement measures. Approximately 9% of the trials were excluded from the

[3] We made no predictions for the comprehension questions as those were solely used to ensure that participants were reading for understanding. We further found no significant difference between participants' mean accuracy responses to the comprehension questions in the experimental conditions, that is as a function of text-belief consistency and mode of presentation.

analyses because of tracking losses, blinks, and incomplete trials (762 out of 8330 trials). Measures of readers' eye movements were computed on the sentence level using Eyelink Analysoija (Hyönä, Kaakinen, & Penttinen, 2016). Eyelink Analysoija uses the fixation report (with default settings) provided by the EyeLink 1000 Data Viewer software to compute the measures. Two first-pass fixation measures as indicators of immediate processes occurring during initial reading of a sentence were computed. The duration of all first fixations for a sentence was computed for the *first-pass reading time*. This measure contained only fixations that landed on an unread word in a given sentence, i.e. all fixations as long as the eyes move forward along a sentence. The *first-pass rereading time* was computed by summing the duration of all reinspective fixations landing on words in a given sentence that had already been read during initial reading. Put differently, fixations that occurred when the eyes returned to already read parts (e.g. words) of a sentence during the first reading of that sentence, were coded as first-pass rereading times. These indicate readers' need to immediately reread a sentence (Kaakinen et al., 2003). In addition, the *number of lookbacks* (i.e., fixations made on previously read text sentences, including backward and forward fixations) was computed as an indicator of delayed processing (for further details on the measures see Hyönä et al., 2003). To control for sentence length, fixation time measures were divided by the number of words per sentence. Lookbacks were coded dichotomously, depending on whether at least one lookback was recorded for a given sentence (yes = 1, no = 0).

We conducted linear mixed model (LMM) analyses for the reading time measures (first-pass reading time and first-pass rereading time) and a generalized linear mixed model (GLMM) analysis with a logit link function for the lookbacks. Subjects and sentences were included as random factors, i.e. the means of subjects as well as sentences were allowed to vary randomly to account for the fact that the factors are sampled from larger random populations. Not only random-intercepts were included, but also by-subject random slopes for belief-consistency, mode of presentation and argument

part to reduce the likelihood of high Type 1 errors rates due to the within-subjects manipulations (see Barr, Levy, Scheepers, & Tily, 2013).

By incorporating crossed random effects for subjects and sentences, the LMM/GLMM analyses can test the hypothesized effects of the independent variables in one single model (for further discussion, see Baayen, Davidson, & Bates, 2008) without the need of running separate by-subjects and by-items analyses. The contrast-coded independent variables text-belief-consistency (-1 = belief-inconsistent, 1 = belief-consistent), mode of text presentation (-1 = blocked, 1 = alternating), argument part (-1 = reason, 1 = claim), and belief-strength (grand-mean centered) and their interactions were included as predictors with fixed effects in the model. In addition, to control for ordering and topic effects as well as for differences in reading ability, the order in which participants read the text (belief-consistent text first vs. belief-inconsistent text first, contrast-coded), the text topic (vaccination or global warming, contrast-coded) and participants' reading skills (grand-mean centered) were incorporated as control predictor. Further, participants received two texts on each text stance for each topic. Given that texts might differ slightly, we controlled for text specific effects by including the text (text 1 or text 2, contrast-coded) as control factors in the analysis. Finally, given that participants came twice to the lab, it might have been that participants processing and comprehension differed if it was their first or second time taking part in the study. Hence, measurement time (first time at the lab, second time at the lab, contrast-coded) was included as control predictor, too.

The lme4 package (Version 1.1-12, Bates et al., 2015) with Restricted Maximum Likelihood Estimation (REML) was used to estimate the models in R (Version 3.2.2., R Core Team, 2014). The lmertest package (Version 2.0-29, Kuznetsova, Brockhoff, & Christensen, 2015) was used for the significance tests of the fixed effects in the linear mixed models. The lmertest package uses the so-called Satterthwaite approximation (Satterthwaite, 1946) for computing the degrees of freedom for the *t*-tests of the fixed effects, which leads to reliable Type-I error rates even in small samples and when

the distribution of random effects is misspecified (Manor & Zucker, 2004). Note that the Satterthwaite approximation may lead to varying degrees of freedom for fixed effects of predictors located on the same level of analysis. We report only significance tests relevant to the hypotheses, that is, main fixed effects of the independent variables and interaction effects with our main predictor belief-consistency. Table 3 provides the first-pass reading times and first-pass rereading times as a function of text-belief consistency, mode of presentation and argument part.

First-pass reading time. In line with earlier eye-movement studies on inconsistency effects (Rinck et al., 2003) and as expected, no effects of text-belief consistency ($t(258) = 0.49$, *n.s.*) and presentation mode ($t(26) = -1.41$, *n.s.*) on first-pass reading time were found. However, first-pass reading time was longer for claims ($M = 246.7$, $SE = 8.9$) than for reasons ($M = 223.9$, $SE = 6.8$), $t(251) = 3.81$, $p < .05$.

First-pass rereading time. Hypothesis 1 postulated that for participants with strong prior beliefs first-pass rereading time for belief-inconsistent information should be longer compared to first-pass rereading time for belief-consistent information. In addition, this effect was supposed to be enhanced in a blocked mode of presentation (Hypothesis 2). We further assumed that first-pass rereading time for belief-inconsistent claims should be more pronounced in a blocked mode of presentation (Hypothesis 3). At last, the blocked mode of presentation was expected to yield longer first-pass rereading time for belief-consistent reasons, whereas the alternating mode of presentation was expected to yield longer first-pass rereading times for belief-inconsistent reasons (Hypothesis 4).

In support of Hypothesis 1, we found an interaction of belief strength and text-belief consistency, $t(530) = -2.01$, $p < .05$. Conditional effects of text-belief consistency and presentation mode were computed for high and low values of the moderating variable belief strength to interpret the interaction (Aiken & West, 1991, Chapter 7). The effects were estimated at conditional values one standard deviation above and below the mean of belief strength. For participants with strong prior beliefs, there

were longer first-pass rereading times for belief-inconsistent information (M = 64.8, SE = 5.6) compared to belief-consistent information (M = 56.7, SE = 5.6), $t(321) = 1.77$, $p = .04$ (one-tailed). For participants with low values of belief strength, there was, however, no difference in first-pass rereading time for belief-inconsistent information (M = 50.3, SE = 5.6) and for belief-consistent information (M = 50.5, SE = 5.5), $t(320) = -0.04$, *n.s.*.

We also found a three-way interaction of mode of text presentation, text-belief consistency and belief strength, $t(543) = 2.33$, $p < .05$ (one-tailed). In support of Hypothesis 2, participants with strong prior beliefs (1 *SD* above the mean) showed longer first-pass rereading times for belief-inconsistent information compared to belief-consistent information in a blocked mode of presentation ($t(520) = 2.67$, $p < .05$). For such participants, however, no difference in first-pass rereading times for belief-consistent and belief-inconsistent information was found when the texts were alternately presented, $t(449) = 0.22$, *n.s.* (Figure 2). The results of participants with weaker beliefs (1 *SD* below the mean) revealed no differences in first-pass rereading times for belief-consistent and belief-inconsistent information regardless of whether the texts were presented in a blocked ($t(415) = -0.50$, *n.s.*) or in an alternating mode, $t(536) = 0.37$, *n.s.*

Moreover, we found a three-way interaction of argument part, mode of text presentation and text-belief consistency, $t(7319) = 2.59$, $p < .05$. As expected, participants' first-pass rereading times for belief-inconsistent claims were longer in a blocked mode of presentation compared to an alternating mode of presentation, $t(236) = 3.70$, $p < .05$. No difference emerged in first-pass rereading times for belief-consistent claims in the alternating compared to the blocked mode of presentation, $t(254) = 1.45$, *n.s*. (see Figure 3). Thus, the pattern of first-pass rereading of claims supported Hypothesis 3. For reasons, we found no difference between first-pass rereading times for belief-consistent and belief-inconsistent reasons in a blocked mode of presentation, $t(256) = 0.72$, *n.s.* Thus, Hypothesis 4a was not supported. However, first-pass rereading times for belief-inconsistent reasons

were longer compared to first-pass rereading times for belief-consistent reasons when the texts were alternately presented, $t(254) = 2.23$, $p < .05$ (see Figure 3), which supported Hypothesis 4b.

We also found main effects for argument part ($t(120) = 2.69$, $p < .05$) and mode of text presentation, $t(43) = -2.83$, $p < .05$. Similar to the results for first-pass reading time, first-pass rereading time was longer for claims ($M = 62.3$, $SE = 6.4$) compared to reasons ($M = 48.8$, $SE = 3.4$). In addition, first-pass rereading time was longer in a blocked ($M = 61.1$, $SE = 5.3$) compared to an alternating text presentation ($M = 50.1$, $SE = 4.4$).

Lookbacks. Overall, participants made few lookbacks during second-pass reading ($M = 0.29$, $SD = 0.46$). Only 2232 out of 7568 sentences with valid data (29.5%) received a lookback. We expected readers to show more lookbacks for belief-inconsistent information in an alternating compared to a blocked mode of presentation (Hypothesis 5). No main effects of text-belief consistency ($z = -1.27$, *n.s.*), argument part ($z = -0.93$, *n.s.*), and mode of presentation ($z = -0.42$, *n.s.*) were found. However, we found a significant interaction of text-belief consistency and mode of text presentation, $z = -3.08$, $p < .05$. In a blocked mode of presentation, belief-consistent and belief-inconsistent sentences had the same likelihood of being refixated, $z = -0.28$, *n.s.* In an alternating mode of presentation, belief-inconsistent sentences were more likely to receive a lookback compared to belief-consistent sentences, $z = 2.57$, $p < .05$. Figure 4 displays the predicted (conditional) probabilities of lookbacks, back-transformed from the logit-link model with estimated standard errors. In sum and in line with Hypothesis 5, participants engaged in delayed reading behavior for belief-inconsistent information when reading texts in an alternating mode of presentation.

Effects of Text-Belief Consistency and Level of Lookbacks on the Essay Task

The essay tasks (one essay task per participant for each topic) were scored for the number of arguments that were consistent with the pro (that is belief-consistent) and the contra (that is the belief-inconsistent) stance of the texts. We used a coding schema that distinguished between the arguments

presented in the texts (sixteen categories per topic—four texts with four arguments each). Participants' essays were scored by two independent raters and agreement between raters was good with an inter-rater reliability (Cohen's κ) of .87 (.96 for distinguishing between pro and contra arguments). Differences in ratings were resolved through discussion.

The data analysis was conducted with a LMM for the arguments generated in the essay task. For each argument, we analyzed whether or not it was stated in the essay. Hence, the possible number of arguments ranged from 0 to 16. Two related aims guided the analysis. First, we were interested in whether the effects of text-belief consistency and mode of text presentation investigated in earlier studies with a recognition task would also occur in the essay task, which could represent an alternative mean for assessing readers' mental model of the scientific controversy. We predicted that participants would generate more belief-consistent arguments overall as a result of the myside bias (Hypothesis 6), but the effect was expected to be reduced in an alternating mode of presentation (Hypothesis 7). In addition, we expected that participants with a higher amount of lookbacks would especially benefit from an alternating mode of text presentation (Hypothesis 8). For the essay data set, belief-consistent and belief-inconsistent arguments were the unit of analyses. Similar to the analyses for the eye measurement data, the specified LMM included the contrast-coded independent variables belief-consistency (-1 = belief-inconsistent (BIC), 1 = belief-consistent (BC)) and mode of presentation (-1 = blocked, 1 = alternating) and their interactions as predictors with fixed effects. In addition, participants' random effects for the number of lookbacks and its interaction with text-belief consistency and mode of text presentation were included as additional predictors to investigate Hypothesis 8. The order of texts (contrast-coded), belief strength (grand-mean centered), reading ability (grand-mean centered) and the measurement time (contrast-coded) were included as control

factors in the analysis. Random-intercepts for subjects and text topic, as well as by-subject random slopes for belief-consistency and mode of presentation were included.

In line with Hypothesis 6, participants generated more belief-consistent arguments overall (M = 2.61, SE = 0.18) than belief-inconsistent arguments (M = 1.59, SE = 0.20), $t(34) = 4.42$, $p < .05$. In addition, a significant interaction of text-belief consistency and mode of text presentation was found, $t(28) = -2.23$ $p < .05$. The pattern of effects underlying the interaction supported Hypothesis 7: In a blocked mode of presentation, participants produced nearly twice as much belief-consistent than belief-inconsistent arguments (BC: M = 2.96, SE = 0.23, BIC: M = 1.56, SE = 0.22), $t(59) = -4.76$, $p < .05$. However, when participants read belief-consistent and belief-inconsistent texts in an alternating mode, the text-belief consistency effect was attenuated but not fully eliminated (BC: M = 2.25, SE = 0.23, BIC: M = 1.62, SE = 0.26), $t(58) = -2.26$, $p < .05$.

In line with Hypothesis 8, we also found that the relationship between text-belief consistency and mode of presentation was moderated by participants' individual level of lookbacks, $t(34) = -2.45$, $p < .05$ (see Figure 5). Participants with a higher level of lookbacks (point estimate at 1 *SD* above the mean) were able to benefit from the alternating mode of text presentation. These participants reported a similar amount of belief-consistent and belief-inconsistent arguments after they read the texts in an alternating mode of presentation, $t(59) = -0.51$, *n.s.* After having read the texts in a blocked mode of presentation, these participants, however, reported more belief-consistent arguments compared to belief-inconsistent arguments, $t(58) = -4.75$, $p < .05$. In contrast, an alternating mode of presentation was not beneficial for participants with a lower level of lookbacks (point estimate at 1 *SD* below the mean), because such participants reported more belief-consistent arguments than belief-inconsistent arguments regardless of the presentation mode (alternating: $t(57) = -2.55$, $p < .05$; blocked: $t(59) = -2.20$, $p < .05$). This pattern of results supported Hypothesis 8.

Discussion

The present study investigated the cognitive processes underlying the text-belief consistency effect in readers' representation of scientific discourse and its moderation by the mode of text presentation. Moreover, the differential functions of scientific argument claims and reasons were assessed and comprehension outcomes were investigated with an essay task.

The results for the essay task fit nicely into existing research investigating how the mode of presentation influences the memory for belief-consistent and belief-inconsistent arguments (e.g., Wiley, 2005) and multiple texts (e.g., Maier & Richter, 2013). As earlier research has shown, a blocked mode of presentation leads to a weaker memory for belief-inconsistent arguments or texts, whereas similar memory levels are achieved for belief-consistent and belief-inconsistent arguments or texts if such arguments or texts are presented alternatingly.

Further, the eye tracking data assessed in our study are able to shed some light on the cognitive processes that underlie the text-belief consistency effect in the memory representation. In line with our assumptions, results revealed differences in participants' immediate and delayed processing of belief-consistent and belief-inconsistent claims and reasons, which varied as a function of presentation mode. First, we found longer first-pass rereading times for belief-inconsistent information in readers with strong prior beliefs. This result is consistent with earlier research showing that readers monitor incoming text information for consistency with prior text information and prior knowledge and that such processes lead to slow-downs in reading (for instance, Albrecht & O'Brien, 1993; Singer, 2006). Our results further extend this research in two ways. First, longer first-pass rereading times for belief-inconsistent information in readers with strong beliefs suggest that readers also monitor text-belief consistency during comprehension if strong beliefs are available. This is consistent with results from Wolfe, Tanner, and Taylor (2013) who also found reading slow-downs for belief-inconsistent sentences. Second and even more important, a difference in first-pass rereading times for belief-

consistent and belief-inconsistent sentences suggests that such a monitoring process already occurs in early and immediate comprehension processing stages. In other words, comprehending controversial texts that are consistent or inconsistent with readers' prior beliefs seems to include an immediate monitoring process that takes the belief-consistency of information into account. However, this immediate monitoring effect is only able to occur if strong prior beliefs are available as epistemic background.

In addition, we found that longer first-pass rereading times occur especially for belief-inconsistent claims when participants read the controversial texts in a blocked mode of presentation. This result is consistent with the notion that belief-consistency is especially activated and used for validation during claim processing (Voss et al., 1993; Wolfe et al., 2009). We think that a blocked mode of text presentation makes it less likely that readers become fully aware of intertextual inconsistencies, which decreases the probability that immediate monitoring of belief-consistency is accompanied by delayed repair processes to resolve such inconsistencies (van den Broek et al., 2015). Even when belief-inconsistent information leads to a disruption in immediate reading, the information cannot be readily integrated into readers' evolving mental model (Schroeder et al., 2008). Our results fit this interpretation, because a blocked mode of presentation led to stronger mental models for the belief-consistent text compared to the mental model for the belief-inconsistent text. These findings can be interpreted in terms of an immediate validation process that prevents the acquisition of potentially unwanted information (Richter, 2015). Such an immediate validation process protects the mental system of the reader in the case of false and incorrect knowledge, but it hinders situation model updating in the case of belief-inconsistent information. In detail, readers' processing is disrupted by belief-inconsistent information, which suggests that belief-consistency is monitored during immediate processing. However, readers often seem to forego further delayed attempts to resolve the inconsistencies, such as making lookbacks to reinstate information in working memory, and simply

continue reading. The lack of attending to belief-inconsistent information in delayed processing fosters a biased mental representation of scientific discourse.

In contrast, an alternating mode of text presentation increased immediate and delayed processing for belief-inconsistent information and was beneficial for comprehension. We found longer first-pass rereading times for belief-inconsistent reasons and a higher likelihood of lookbacks to belief-inconsistent information in an alternating text presentation. Moreover, such a mode of presentation advanced processing belief-inconsistent reasons which are crucial for understanding and critically evaluating an argument (Wolfe et al., 2009). Together with the attenuating effect of an alternating presentation mode on the text-belief consistency effect, especially for participants with a higher amount of lookbacks, these results support the assumption that an alternating mode of presentation can reduce the text-belief consistency effect.

One likely interpretation of the findings is that an alternating mode of presentation increased participants' awareness of inconsistencies and their need for repair processes. This interpretation coheres well with earlier research that shows that more delayed initiated lookbacks might indicate a kind of repair process during which the reader tries to resolve an inconsistency, but only if the inconsistency is consciously noted (see Rinck et al., 2003). Given that mode of presentation was varied as a within participant factor, such an improved detection of intertextual inconsistencies by an alternating mode of presentation might have increased participant's standards of coherence (van den Broek et al., 2015). In a given reading situation, readers follow specific standards of what constitutes (sufficiently) good comprehension. These standards influence the extent that readers engage in strategic processing during comprehension. An alternating mode of text presentation might have helped readers to become aware of inconsistencies between the texts, which probably prompts them to regulate their comprehension processes to restore coherence by resolving the inconsistency. A similar process might occur during reading refutational texts in which the coactivation of false and accurate

beliefs is able to foster the integration of newly encoded information with previous knowledge and beliefs (Kendeou & O'Brien, 2014).

Our method contributes to an understanding of how prior beliefs affect comprehension of scientific texts in a particular reading situation. In detail, given that presentation mode was varied within participants in our experimental design, we are able to make some suggestions with regard to the questions under which circumstances belief-consistent or belief-inconsistent information receives more attention and when processes of belief protection or belief reflection prevail. Our results suggest that if the detection of inconsistencies between texts is unlikely in a reading situation and the amount of delayed processing is as a result low, belief-consistent information will predominate processing and the representation of the scientific controversy. One interpretation of the pattern of results we found is that immediate validation based on prior beliefs seems to involve a belief-protection mechanism (McCrudden & Sparks, 2014) that aims to avoid information processing that might challenge a reader's point of view (similar to selective exposure, Festinger, 1957). Put differently, if the amount of delayed processing is low, more attention is paid to belief-consistent information. In this type of processing, readers are able to construct a good mental model for belief-consistent arguments but fail to build a solid understanding of belief-inconsistent arguments as indicated by the results of the essay task. However, if readers become consciously aware of inconsistencies between conflicting texts as a result of an alternating mode of text presentation, they are more likely to follow a belief-reflection purpose during reading (McCrudden & Sparks, 2014). This type of processing seems to include an emphasis on belief-inconsistent information in immediate and delayed processing. Hence more attention is paid to belief-inconsistent information. This leads to a better understanding of both argumentative sides in a scientific controversy and a stronger likelihood to resolve or explain the inconsistencies (Edwards & Smith, 1996; Richter, 2015; van den Broek et al., 2015).

Our interpretation of the results is consistent with standards of relevance theory (e.g., McCrudden et al., 2011) and perspective effects (e.g., Anderson & Pichert, 1978), which suggest that especially textual information relevant for a reading purpose or perspective receives more attention during reading and is more strongly integrated in the reader's mental model of the text (e.g., Kaakinen & Hyönа, 2005, 2011). Readers' prior beliefs might also serve as standards of relevance in the manner described above and thus might determine the instrumental value of a text segment in relation to readers' reading task. However, one important difference between relevance and prior belief effects is that relevance instructions are given directly before reading. In contrast, prior beliefs are activated during reading, providing readers with less opportunity to actively control the influence of their beliefs.

One limitation of our study is that the interpretation of the immediate processing indicator is somewhat ambiguous, because longer first-pass rereading times can indicate enhanced integration processes or processing difficulties (i.e., disruptions). From our point of view, the assumption that longer first-pass rereading times for claims indicate reading disruptions as a result of inconsistencies between the argumentative stance of the claim and one's own argumentative stance in the scientific controversy is plausible. Argumentative reasons, however, are not directly consistent or inconsistent with readers' prior beliefs, but they more or less support the claim provided in the text. Thus, we consider it more reasonable to assume that longer first-pass rereading times of reasons indicate enhanced integration. Further research using event-related potentials (ERP) as an indicator of immediate processing could test this hypothesis. ERP allows for a dissociation between detection of incoherence and integration of incoming information in early and non-strategic stages of comprehension by using time-frequency analysis (see Steele, Bernat, van den Broek, Collins, Patrick, & Marsolek, 2013). Time-frequency analysis of ERPs from reading belief-consistent and belief-inconsistent claims and reasons might provide a means to assess whether longer first-pass rereading

times for belief-inconsistent claims indicate comprehension difficulties as a result of inconsistency detection and whether longer first-pass rereading times for belief-consistent reasons indicate enhanced integration of this type of information.

A second potential limitation of our study is that we used an essay task to assess the extent to that participants constructed a balanced mental model of the controversy that includes arguments from both sides of the issue. Such tasks are often used for this purpose (e.g., Anmarkrud et al., 2013; Richter, 2003; see Richter & Maier, 2017), and they seem to be a natural way to assess understanding of controversial, argumentative texts. It must be noted, however, that the essay task also explicitly prompted participants to express their own beliefs, which might have increased a text-belief consistency effect. At any rate, the essay task used in the present study differs in several respects from standard memory and comprehension tasks used in the earlier studies (e.g. Maier & Richter, 2013; Wiley, 2005). However, a recent overview from Richter and Maier (2017) discussing text-belief consistency effects for different outcomes of multiple text comprehension, suggests that the text-belief consistency effect generalizes across different types of tasks that involve readers' beliefs to different degrees.

Compared to earlier work on belief effects on memory (e.g., Wiley, 2005) and relevance effects on eye movements (e.g., Kaakinen, & Hyönä, 2005, 2011), our study did not use a balanced sample with equal numbers of participants who prefer each argumentative position on the two issues. The reason why we chose this design is that only a between-text variation of text-belief consistency seemed suitable to ensure a strong manipulation of this variable. However, despite the fact that material effects cannot be ruled out completely, all efforts were made to keep the texts and the arguments within the texts as parallel as possible and t-tests for paired samples did not reveal any differences between the texts. Moreover, the effect of primary theoretical interest was the interaction effect of text-belief consistency with mode of presentation (blocked vs. alternating), and presentation mode was varied

experimentally as a within-participant factor. Nevertheless, future studies should be conducted with balanced samples, too.

With the growing amount of information available in the World Wide Web and the ease in which it can be accessed, encountering belief-inconsistent information is a fairly common occurrence. The access to various stances on a particular topic increases the relevance to effectively process and comprehend the information. However, the results of the present study suggest that readers have difficulties in processing and comprehending belief-inconsistent arguments and that the difficulties are partly due to the effects of an immediate validation process included in comprehension. An alternating mode of presenting belief-consistent and belief-inconsistent texts can foster the processing and comprehension of belief-inconsistent information, but the effectiveness of an alternating mode of presentation strongly depends on the extent of delayed reading behavior that readers employ. We believe that this study contributes to advancing fruitful hypotheses for investigating other potential means that might reduce the text-belief consistency effect and methods for training readers how to process belief-inconsistent information in the comprehension of multiple science-related texts in an effective manner.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 106l–1070.

Amstad, T. (1978). *Wie verständlich sind unsere Zeitungen?* [How understandable are our newspapers?]. Unpublished doctoral dissertation, University of Zürich, Switzerland.

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior, 17*, 1–12.

Anmarkrud, Ø., McCrudden, M. T., Bråten, I., & Strømsø, H. I. (2013). Task-oriented reading of multiple documents: Online comprehension processes and offline products. *Instructional Science*, *41*, 873–894.

Ariasi, N., & Mason, L. (2011). Uncovering the effect of text structure in learning from a science text: An eye-tracking study. *Instructional Science, 39*, 581–601.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Sigmann, H., & Grothendieck, G. (2015). *Lme4: linear mixed-effects models using Eigen and S4* [Software]. R-package version 1.1-10. Retrieved November 1, 2015 from:http://cran.r-project.org/package=lme4

Braasch, J. L. G., Goldman, S. R., & Wiley, J. (2013). The influences of text and reader characteristics on learning from refutations in science texts. *Journal of Educational Psychology*, *105*, 561–578.

Britt, M. A., Richter, T., & Rouet, J. -F. (2014). Scientific Literacy: The role of goal-directed reading and evaluation in understanding scientific information. *Educational Psychologist, 49*, 104–122.

Burton, C., & Daneman, M. (2007). Compensating for a limited working memory capacity during reading: Evidence from eye movements. *Reading Psychology, 28,* 163–186.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science education. *Review of Educational Research, 63*, 1–49.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Eagly, A. H., Kulesa, P., Brannon, L. A., Shaw, K., & Hutson-Comeaux, S. (2000). Why counterattitudinal messages are as memorable as proattitudinal messages: The importance of active defense against attack. *Personality and Social Psychology Bulletin, 26*, 1392–1408.

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology, 71*, 5–24.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Hart, W., Albarracin, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill. L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin, 135*, 555–588.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70.

Hyönä, J., Kaakinen, J. K., & Penttinen, M. (2016). *Eyelink Analysoija - Eye movement data analysis software for computing sentence-level measures* [Computer program and manual]. Unpublished computer program. Retrieved September 15, 2016 from https://seafile.utu.fi/d/18cd89253d /

Hyönä, J., Lorch, R. F. Jr., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Amsterdam, The Netherlands: Elsevier.

Isberner, M. -B., & Richter, T. (2014). Comprehension and validation: Separable stages of information processing? A case for epistemic monitoring in language comprehension. In D. N. Rapp & J. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 245–276). Boston, MA: MIT Press.

Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When discredited information in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1420–1436.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, *80*, 557–571.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Kaakinen, J. K., & Hyönä, J. (2005). Perspective effects on expository text comprehension: Evidence from think-aloud protocols, eyetracking, and recall. *Discourse Processes, 40*, 239–257.

Kaakinen, J. K., & Hyönä, J. (2008). Perspective-driven text comprehension. *Applied Cognitive Psychology, 22*, 319-334.

Kaakinen, J. K., & Hyönä, J. (2011). Online processing of and memory for perspective relevant and irrelevant text information. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 223–242). Greenwich, CT: Information Age Publishing.

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2003). How prior knowledge, working memory capacity, and relevance of information affect eye-fixations in expository text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 447–457.

Kardash, C. A. M., & Scholes, R. J. (1996). Effects of pre-existing beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology, 88*, 260–271.

Kendeou, P., & O'Brien, E. J. (2014). The Knowledge Revision Components (KReC) framework: Processes and mechanisms. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 353–377). Cambridge, MA: MIT Press.

Kendeou, P., Walsh, E., Smith, E. R., & O'Brien, E. J. (2014). Knowledge revision processes in refutation texts. *Discourse Processes, 51,* 374–397.

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

Knobloch-Westerwick, S., & Meng, J. (2011). Reinforcement of the political self through selective exposure to political messages. *Journal of Communication, 61*, 349–368.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R. H. B. (2015). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)*. R-package version 2.0-29. Retrieved November 1, 2015 from: http://cran.r-project.org/web/packages/lmerTest/index.html

Limon, M., & Mason, L. (2002). *Reconsidering conceptual change: Issues in theory and practice*. Dordrecht, The Netherlands: Kluwer.

Maier, J., & Richter, T. (2013). Text-belief consistency effects in the comprehension of multiple texts with conflicting information. *Cognition and Instruction, 31*, 151–175.

Maier, J., & Richter, T. (2014). Fostering multiple text comprehension: How metacognitive strategies and motivation moderate the text-belief consistency effect. *Metacognition & Learning, 9*, 54–71.

Maier, J., & Richter, T. (2016). Effects of text-belief consistency and reading task on the strategic validation of multiple texts. *European Journal of the Psychology of Education, 31*, 479–497.

Manor, O., & Zucker, D. M. (2004). Small sample inference for the fixed effects in the mixed linear model. *Computational Statistics & Data Analysis, 46*, 801–817.

McCrudden, M. T., & Sparks, P. C. (2014). Exploring the effect of task instructions on topic beliefs and topic belief justifications: A mixed methods study. *Contemporary Educational Psychology, 39*, 1–11.

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals, and text processing: A mixed methods study. *Contemporary Educational Psychology, 35* (4), 229–241.

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2011). Relevance in text comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 1-17). Greenwich, CT: Information Age Publishing.

O'Brien, E. J., & Cook, A. E. (2015). Models of discourse comprehension. In A. Pollatsek & R. Treiman (Eds.), *Handbook on reading* (pp. 217–231). New York: Oxford University Press.

O'Brien, E. J., & Myers, J. L. (1999). Text comprehension: A view from the bottom up. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 35–54). Mahwah, NJ: Erlbaum.

R Core Team (2014). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/

Richter, T. (2003). *Epistemologische Einschätzungen beim Textverstehen* [Epistemic validation in text comprehension]. Lengerich: Pabst.

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes*, *52*, 337–355.

Richter, T. & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A Two-step Model of Validation. *Educational Psychologist , 52*, 148-166.

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology, 96*, 538–598.

Richter, T., & van Holt, N. (2005). ELVES: Ein computergestütztes Diagnostikum zur Erfassung der Effizienz von Teilprozessen des Leseverstehens. *Diagnostica, 51*, 169–182.

Rinck, M., Gámez, E., Díaz, J. M., & de Vega, M. (2003). Processing of temporal information: Evidence from eye movements. *Memory & Cognition, 31*, 77–86.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32*, 880–892.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrika Bulletin, 2*, 110–114.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. *Journal of Memory and Language, 25*, 279–294.

Schroeder, S., Richter, T., & Hoever, I. (2008). Getting a picture that is both accurate and stable: Situation models and epistemic validation. *Journal of Memory and Language, 59*, 237–259.

Singer, M. (2006). Verification of text ideas during reading. *Journal of Memory and Language, 54*, 574–591.

Steele, V. R., Bernat, E.M., van den Broek, P., Collins, P. F., Patrick, C. J., & Marsolek, C. J. (2013). Separable processes before, during, and after the N400 elicited by previously inferred and new information: Evidence from time–frequency decompositions. *Brain Research*, *1492,* 92–107.

Toplak, M. E., & Stanovich, K. E. (2003). Associations between myside bias on an informal reasoning task and amount of post-secondary education. *Applied Cognitive Psychology, 17*, 851–860.

Toulmin, S. E. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

van den Broek, P., Beker, K., & Oudega, M. (2015). Inference generation in text comprehension: Automatic and strategic processes in the construction of a mental representation. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 94–121). Cambridge, UK: Cambridge University Press.

van Strien, J. L. H., Brand-Gruwel, S., & Boshuizen, H. P. A. (2014). Dealing with conflicting information from multiple nonlinear texts: Effects of prior attitudes. *Computers in Human Behavior, 32*, 101–111.

von der Mühlen, S., Richter, T., Schmid, S., Schmidt, L. M., & Berthold, K. (2016). Judging the plausibility of argumentative statements in scientific texts: A student-scientist comparison. *Thinking and Reasoning, 22*, 221–246.

Voss, J. F., Fincher-Kiefer, R., Wiley, J., & Silfies, L. N. (1993). On the processing of arguments. *Argumentation, 7,* 165–181.

Walczyk, J. J., & Taylor, R. W. (1996). How do the efficiencies of reading subcomponents relate to looking back in text? *Journal of Educational Psychology, 88,* 537–545.

Wiley, J. (2005). A fair and balanced look at the news: What affects memory for controversial arguments? *Journal of Memory and Language, 53*, 95–109.

Wolfe, C. R., Britt, M. A., & Butler, J. A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication, 26*, 183–209.

Wolfe, M. B., Tanner, S. M., & Taylor, A. (2013). Processing and representation of arguments in one-sided texts about disputed topics. *Discourse Processes, 50,* 457–497.

Table 1

Text Characteristics of the Eight Experimental Texts as a Function of Text Topic and Text-belief Consistency

Text Topic	Argumentative Stance	Length[a]	Mean words per sentence	Mean syllables per word	Reada-bility[b]	Plausibility Scale[c] *M* (SE_M)	Understandability Scale[c] *M* (SE_M)	Number of Arguments[c] *M* (SE_M)	Interesting-ness[c] *M* (SE_M)
Global warming	BIC 1	594	14.49	2.12	42	4.10 (0.27)	4.43 (0.21)	4.44 (0.50)	4.00 (0.34)
	BIC 2	584	14.60	2.04	46	3.74 (0.17)	4.38 (0.17)	4.13 (0.47)	3.50 (0.27)
	BC 1	680	15.81	2.03	46	4.62 (0.31)	4.40 (0.30)	3.55 (0.25)	3.55 (0.51)
	BC 2	610	14.52	2.14	40	4.86 (0.22)	4.51 (0.17)	4.27 (0.27)	4.18 (0.33)
Vaccination	BIC 1	571	14.28	2.05	46	4.33 (0.20)	4.56 (0.20)	3.89 (0.35)	4.22 (0.49)
	BIC 2	621	14.79	2.08	44	4.64 (0.37)	4.76 (0.34)	4.57 (0.43)	4.57 (0.37)
	BC 1	589	14.02	2.11	43	3.79 (0.53)	4.61 (0.32)	3.25 (0.56)	3.63 (0.65)
	BC 2	600	14.36	2.02	47	4.85 (0.23)	4.97 (0.15)	4.00 (0.56)	3.80 (0.25)

Note. BC: belief-consistent; BIC: belief-inconsistent.

[a]Number of words per text. [b]Determined with the German adaption of the Flesch`s Reading Ease Index (Amstad, 1978). [c]Results of the pilot-testing with ratings of 45 university students (response categories ranging from 0 = *not at all* to 6 = *totally;* The plausibility scale consists of six items (Cronbach's α = .74) and the understandability scale consists of nine items (Cronbach's α = .86). Each entry represents the average judgments across all participants.

Table 2

Number of Sentences (Trials) within the 2x2x2 Design (Topic x Text Stance x Text) for Eye-Movement Measures.

Topic	Text Stance	Sentences for Claim and Reasons
Global Warming	Belief-consistent 1	30
	Belief-consistent 2	31
	Belief-inconsistent 1	30
	Belief-inconsistent 2	31
Vaccination	Belief-consistent 1	31
	Belief-consistent 2	31
	Belief-inconsistent 1	30
	Belief-inconsistent 2	31

Table 3

First-pass reading time and first-pass rereading time as Function of Text-belief Consistency, Mode of Presentation and Argument Part

		First-pass reading time				First-pass rereading time			
		BC		BIC		BC		BIC	
Mode of Presentation	Argument Part	*M*	SE_M	*M*	SE_M	*M*	SE_M	*M*	SE_M
Alternating	Claim	249.76	10.67	235.66	10.91	56.69	7.78	52.47	7.68
	Reason	217.45	7.34	223.59	7.68	41.84	3.61	49.25	3.64
	Total	233.60	8.11	229.62	8.42	49.26	4.99	50.86	4.91
Blocked	Claim	256.29	11.05	244.96	10.80	65.03	8.24	75.19	8.23
	Reason	223.54	7.89	231.19	7.58	50.83	4.59	53.26	4.84
	Total	239.91	8.60	238.07	8.30	57.93	5.69	64.22	5.79
Total	Claim	253.02	10.38	240.31	10.38	60.86	7.39	63.83	7.35
	Reason	220.49	7.16	227.39	7.18	46.33	3.67	51.26	3.83
	Total	236.80	7.90	233.85	7.90	53.60	4.89	57.54	4.91

Note. BC: belief-consistent; BIC: belief-inconsistent.

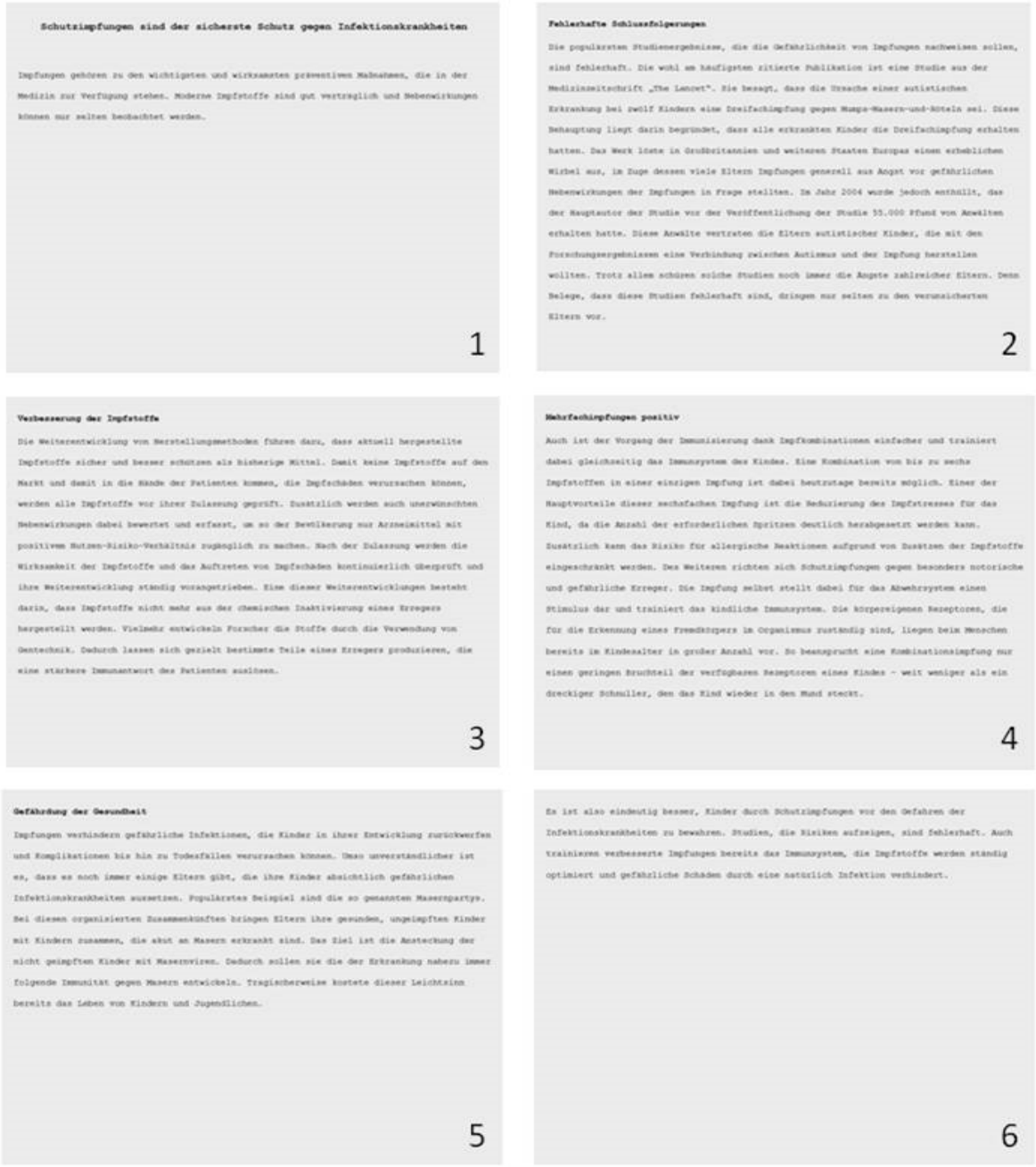

Figure 1. Example of the experimental text in its original formatting (ordering numbers were not provided in the study).

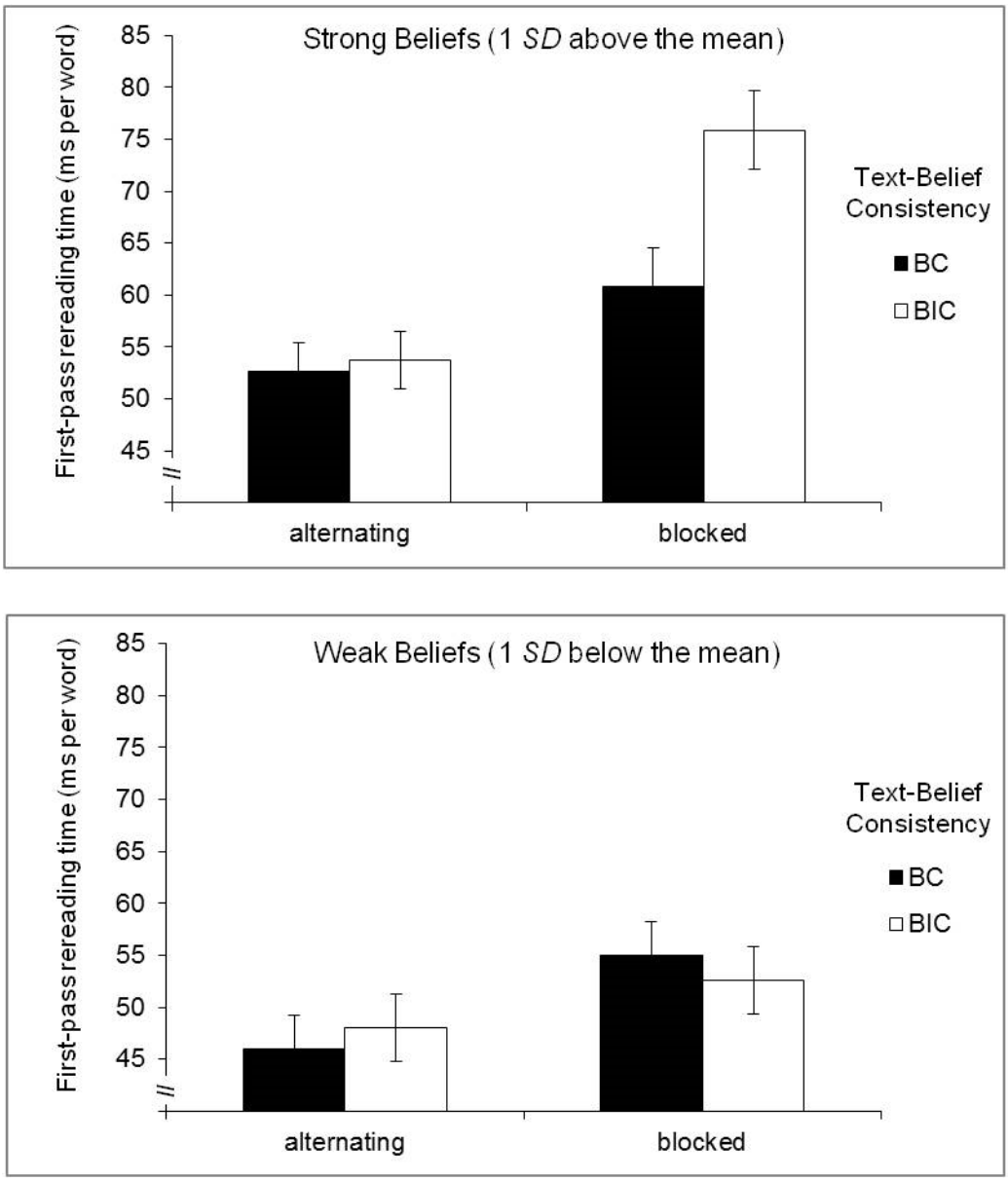

Figure 2. Interaction of beliefs strength, text-belief consistency (BC: belief-consistent vs. BIC: belief-inconsistent) and mode of presentation (alternating vs. blocked) for first-pass rereading times (in ms per word). The effects of text-belief consistency and mode of presentation were estimated at conditional values of one standard deviation above (high level of lookbacks) and one standard deviation below (low level of lookbacks) the mean of belief strength. Error bars represent the standard error of the mean.

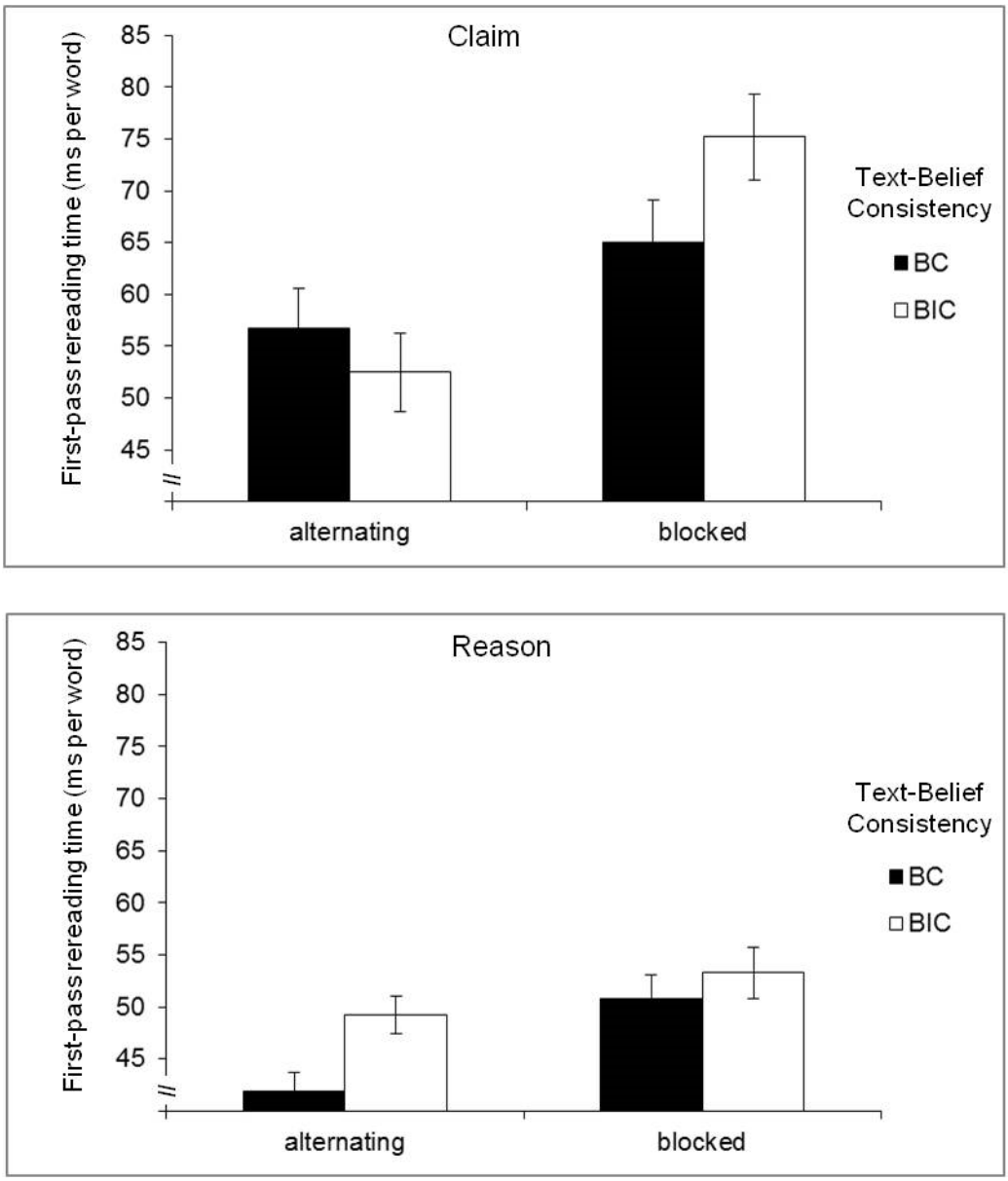

Figure 3. Interaction of argument part (claim vs. reason), text-belief consistency (BC: belief-consistent vs. BIC: belief-inconsistent) and mode of presentation (alternating vs. blocked) for first-pass rereading times (in ms per word). Error bars represent the standard error of the mean.

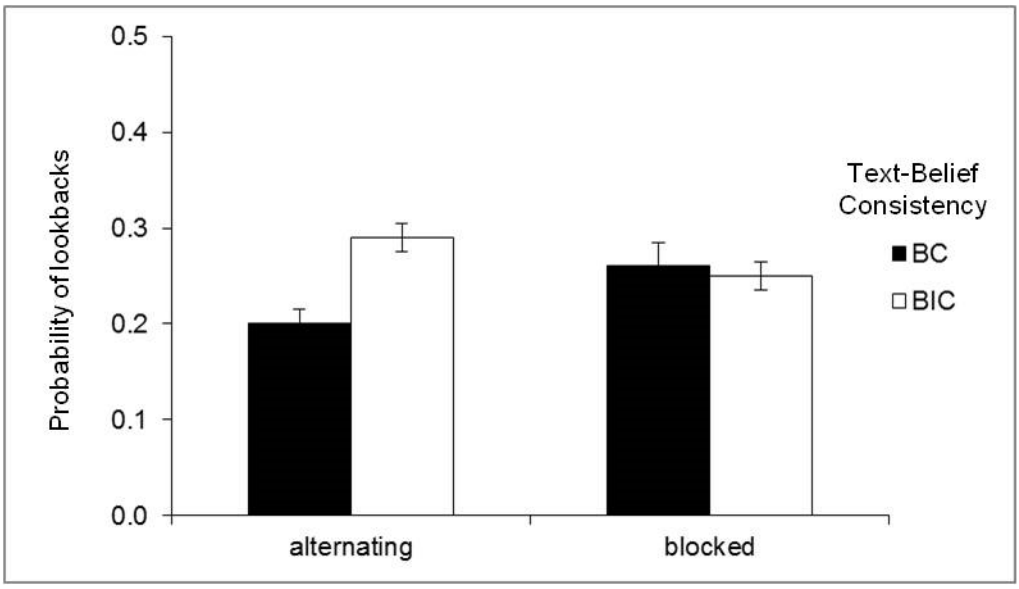

Figure 4. Interaction of text-belief consistency (BC: belief-consistent vs. BIC: belief-inconsistent) and mode of presentation (alternating vs. blocked) for frequency of lookbacks (back-transformed probabilities, error bars represent the standard error of the mean).

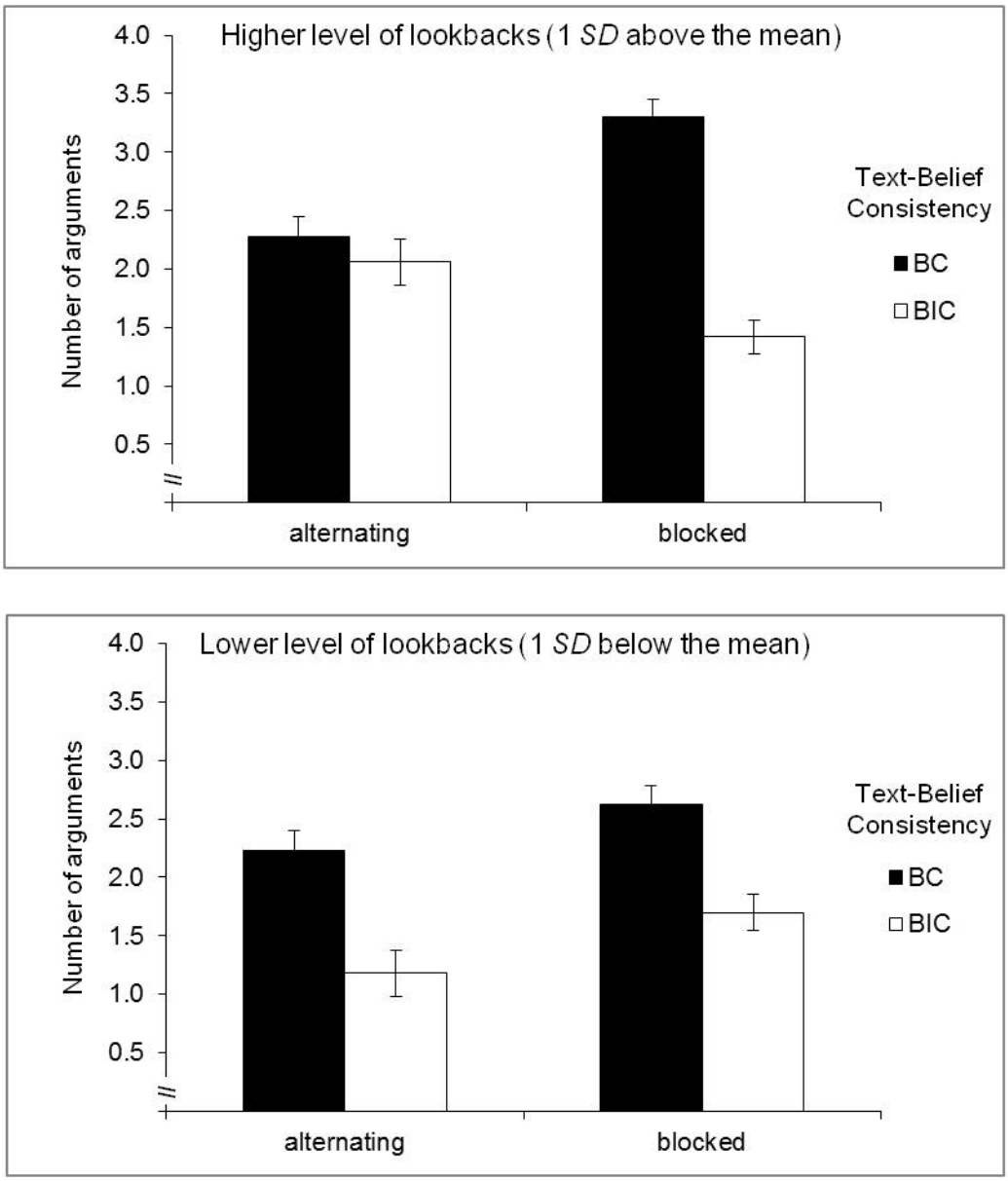

Figure 5. Interaction of text-belief consistency (BC: belief-consistent vs. BIC: belief-inconsistent), mode of presentation (alternating vs. blocked), and participants' individual level of lookbacks for number of arguments generated in the essay task. The effects of text-belief consistency and mode of presentation were estimated at conditional values of one standard deviation above (high level of lookbacks) and one standard deviation below (low level of lookbacks) the mean of lookbacks. Error bars represent the standard error of the mean.