# Fostering Multiple Text Comprehension: How Metacognitive Strategies and Motivation Moderate the Text-Belief Consistency Effect

**Accepted for publication in the journal *Metacognition & Learning* (2013)**

Johanna Maier and Tobias Richter

University of Kassel, Germany

Corresponding author:

Johanna Maier

University of Kassel, Department of Psychology

Holländische Str. 36-38

34109 Kassel, Germany

E-mail: johanna.maier@uni-kassel.de

Telephone: +49-561 804 3582

Fax: +49-561 804 3586

**Acknowledgments**

The research reported in this article was supported by the German Research Association (Deutsche Forschungsgemeinschaft, DFG, grant RI 1100/5-1 and grant RI 1100/5-2). We would like to thank Irina Fix, Sven Gruß, Markus Kumsteller, Lina Krüger, Sascha Müller and Veronika Schmitt for their help in preparing stimulus materials and collecting data.

## Abstract

Learners often have difficulties comprehending multiple texts about controversial scientific issues. In particular, learners with strong prior beliefs tend to construct a one-sided mental representation that is biased towards belief-consistent information (text-belief consistency effect). In the present study we examined the effectiveness of information of three metacognitive strategies tailored to strengthen the comprehension of belief-inconsistent information during multiple text comprehension. According to theories of self-regulated learning, knowledge about relevant metacognitive strategies improves comprehension only when learners are also motivated to use these strategies. These hypotheses were investigated in an experiment in which 85 participants read one belief-consistent and one belief-inconsistent text about a controversial scientific issue. Participants either received information about three metacognitive strategies or no additional information. In addition, participants' motivation was manipulated by providing them with either negative or positive performance feedback or no feedback. As predicted, a text-belief consistency effect was found, which was eliminated by strengthening the situation model for the belief-inconsistent text only when learners received information about relevant metacognitive strategies and were motivated to use these strategies after positive performance feedback.

*Keywords*: beliefs, metacognition, multiple text comprehension, motivation, validation

# Fostering Multiple Text Comprehension: How Metacognitive Strategies and Motivation Moderate the Text-Belief Consistency Effect

When learners use the World Wide Web to inform themselves about currently debated scientific topics, they normally read more than one text. Comprehending multiple texts from the Web not only requires learners to comprehend information from each text but also to integrate the contents of different, sometimes contradictory texts into one coherent referential representation of the scientific topic (Perfetti, Rouet, & Britt, 1999). Hence, learners also need to evaluate the plausibility of information when they use the World Wide Web to acquire new knowledge about a scientific topic (Mason, Boldrin, & Ariasi, 2010). Prior research indicates that learners often have great difficulties in understanding controversial topics on the basis of multiple texts (Rouet, 2006). In particular, they frequently fail to consider and integrate alternative explanations and interpretations into their mental representation of the situation (Britt, Perfetti, Sandak, & Rouet, 1999). One aspect of this problem may be that learners often possess prior beliefs that are closer to one argumentative position in a controversy and that memory is biased toward belief-consistent information (*text-belief consistency effect*, e.g., Eagly & Chaiken, 1993; Maier & Richter, 2013a; Wiley, 2005). The present research examined whether the comprehension of multiple texts on conflicting issues can be improved by attenuating the text-belief consistency effect. In other words, we attempted to foster a more balanced mental representation of a scientific issue by providing information about metacognitive strategies directed at a rational evaluation of the plausibility of information (Mason et al., 2010). Previous research has shown that training of metacognitive strategies can be a powerful means to foster comprehension of multiple texts (Stadtler & Bromme, 2007). Taking a more specific focus, the present study tested the combined effects of information about relevant metacognitive strategies for comprehending belief-inconsistent multiple texts in combination with learners' motivation to use these strategies.

**Multiple Text Comprehension and Prior Beliefs**

Comprehending multiple texts requires that learners not only accumulate facts from texts, but also relate and integrate conflicting arguments from several texts. For example, a student interested in the risks and benefits of vaccinations might encounter texts arguing that vaccinations are the only means to prevent the spread of infectious diseases. The student might later read a contrasting view that the use of polyvalent vaccines overstrains the infantile immune system and some children have died after getting vaccinated. Given the plethora of information on the World Wide Web and the ease to obtain conflicting views on a wide range of topics, it is an important question under what conditions learning with multiple texts and conflicting information is successful (Perfetti et al., 1999). In general, successfully comprehending multiple texts and, more specifically, creating a balanced mental representation of the discussed issue is a challenging task and demands cognitive, metacognitive and motivational resources that learners frequently are not able or willing to invest (Britt et al., 1999; Rouet, 2006).

Learners, however, in many instances seem to rely on a more superficial processing guided by their beliefs. Learners often hold fast to their beliefs even when they are confronted with new information that explicitly corrects or discredits them (Chinn & Brewer, 1993; Johnson & Seifert, 1994; Kardash & Scholes, 1995; Limon & Mason, 2002; Ross, Lepper, & Hubbard, 1975; Vosniadou, 1994; Wiley, 2005). For example, Wiley (2005) asked learners to read a text about a controversial topic (e.g., whether or not abortion should be legal in the U.S.) that included short arguments which were either consistent or inconsistent with learners' prior beliefs. After reading the text, learners were able to recall more belief-consistent than belief-inconsistent arguments. Thus, evidence has shown that not only the evaluation but also the comprehension and memory for conflicting information is subject to some type of myside bias (confirmation bias, Nickerson, 1998) which favors belief-consistent over belief-inconsistent information.

Schema theory provides a framework that can be used to explain the advantage of belief-consistent information in memory construction. According to schema theory, prior beliefs can be construed as a knowledge-based structure, creating specific expectations which then guide the encoding, the interpretation and also the retrieval of information (Pratkanis, 1989). Accordingly, prior beliefs can be seen as a conceptual filter that governs the perception and memory of belief-consistent information (i.e., schema-relevant information). Moreover, schema theory predicts that prior beliefs can bias learners' memory in particular when prior knowledge plays a strong role in the processing of incoming information (Bartlett, 1932). For that reason, the advantage of belief-consistent information is likely to be tied to a higher amount of knowledge-based processing during comprehension.

Such processes are believed to occur during the construction of a *situation model* (van Dijk & Kintsch, 1983). The situation model refers to learners' understanding and their construction of a referential representation of the state of affairs described in a text. Thus, the situation model goes beyond memory for information explicitly stated in the text (i.e., the *propositional text base*) because it involves the integration of new information into learners' knowledge base and the production of inferences.

The hypothesis of a text-belief consistency effect in the situation model construction during multiple text comprehension was supported in a recent experiment by Maier and Richter (2013a). In this experiment, university students read two belief-consistent and two belief-inconsistent texts about a currently debated scientific topic (e.g., the causes of global warming) in different presentation orders. After reading the text, they worked on a recognition task with inference items that were used to assess the strength of the situation model and with paraphrase items that were used to assess memory for text. Results revealed that the situation model for texts consistent with participants' prior beliefs was stronger compared to the situation model for the texts that were inconsistent with participants' prior beliefs. In contrast, the authors found a reverse effect for memory about information explicitly provided by the text (i.e., the propositional text base, Kintsch, 1988). Learners had a better memory for texts that were inconsistent with their prior beliefs in contrast to the texts that were consistent with their beliefs. This finding is in line with the schema-pointer-plus-tag model, which assumes that schema-incongruent information receives a prominent (*tagged*) status in the memory representation of a text (Graesser, 1981; Schank & Abelson, 1977).

However, a text-belief consistency effect in situation model construction might not only result from schema-driven encoding and retrieval processes. Alternatively, learners might also use prior knowledge and beliefs to validate the plausibility of incoming text information (*epistemic validation*, Richter, 2011). The term plausibility has been defined in the literature as the "relative potential truthfulness of incoming information compared to our existing mental representations" (Lombardi, 2012, p. 3, see also Isberner & Richter, in press). Hence, plausibility can be understood as an individual's subjective probability judgment that a piece of information is true. Research suggests further that the validation of text information is an integral part of text comprehension (e.g., Isberner & Richter, 2013; Richter, Schroeder, & Wöhrmann, 2009; Singer, 2006) and is strongly linked to the construction of a situation model (Schroeder, Richter, & Hoever, 2008; Singer, 2006). For example, information judged as implausible has a smaller likelihood of being integrated into learners' situation model (Schroeder et al., 2008; see also Lombardi, Sinatra, & Nussbaum, 2013; Maier & Richter, 2013b). Provided that prior beliefs are easily accessible during reading, a text-belief consistency effect in the situation model construction may also result from fast and efficient validation processes during reading.

In sum, both schema theory and research on epistemic validation may contribute to an explanation of the finding that when learning with multiple texts, prior beliefs often lead to a biased, one-sided mental representation. This representation may be coherent but only provides half of the story about a controversial topic. Hence, reducing the text-belief consistency effect in multiple text comprehension is important in order to ensure that learners understand arguments and evidence that are consistent and that run contrary to their beliefs. The present research addressed the question of how learners can be assisted in using their prior knowledge and beliefs optimally such that a text-belief consistency effect can be avoided.

Previous training studies have revealed that the training of sourcing strategies (paying attention to source information and using it for evaluation and interpretation of text information) while reading multiple historic texts (Britt & Aglinskas, 2002; Nokes, Dole, & Hacker, 2007) and the prompting of metacognitive learning strategies (Stadtler & Bromme, 2007) can improve multiple text comprehension. Moreover, evidence from correlational studies has shown that sourcing strategies benefit the comprehension of multiple documents in history (Wineburg, 1991) as well as science-related texts (Strømsø, Bråten, & Britt, 2010). In contrast to these studies, the focus of the present study was not on source evaluation but on processes of *content evaluation*. To be precise, we investigated metacognitive strategies that ensure the strengthening of belief-inconsistent information comprehension.

## Metacognition and (Multiple) Text Comprehension

Learners' multiple text comprehension benefits from knowledge of the cognitive processes involved in comprehension and the ability to control the direction, intensity and persistence of these cognitive processes. The term metacognition refers to both of these aspects of skilled information processing—the knowledge about cognition and the regulation of cognition (Baker & Brown, 1984; Baker & Beall, 2009; Flavell, 1976).

*Comprehension monitoring* refers to the control component of metacognition that combines evaluation and regulation of comprehension (e.g., Baker, 1985, 1989; Myers & Paris, 1978). For example, Baker (1985) assumed that comprehension problems, such as an awareness of an inconsistency between two text passages, can only be detected by learners when they evaluate their performance with regard to specific evaluation standards. Baker (1985) identified seven different evaluation standards that refer to lexical, syntactic and semantic categories. Two of these standards, the internal and external consistency standards of the higher-level semantic category, are likely to play a critical role for processing belief-inconsistent information. The *external consistency standard* focuses on the detection of inconsistencies between text information and learners' prior knowledge whereas the *internal consistency standard* is directed at inconsistencies between ideas in a text. Baker (1985) used the error-detection paradigm with expository texts that contained nonsense word (to detect applications of the lexical standard), prior knowledge violations (to detect applications of the external consistency standard), or textual inconsistencies (to detect violations of the internal consistency standard). Using this method, Baker (1985) found that 95 percent of the student participants applied the lexical standard, whereas the external and the internal consistency standards were applied infrequently unless participants were explicitly instructed to use them. However, using the external and internal consistency standards was associated with better comprehension. Similarly, Stadtler and Bromme (2004) found that lay people often failed to use metacognitive strategies (as revealed by think-aloud protocols) while searching the World Wide Web, but prompting learners to use metacognitive strategies in comprehending multiple texts on the World Wide Web supported their learning processes and knowledge acquisition (Stadtler & Bromme, 2007).

The results are noteworthy compared with studies assessing learners' comprehension monitoring during reading with more indirect measures such as reading times or eye-tracking. These studies suggest that learners—even when they fail to verbally report inconsistencies—seem to still encounter comprehension problems (e.g., Baker & Anderson, 1982; see also the reading time studies reviewed by Baker, 1989; Singer, 2006). Accordingly, comprehension problems appear not to be the result of inconsistency detection failures (comprehension monitoring) but instead are due to insufficient regulation of comprehension after noticing that coherence formation is disrupted.

**Metacognitive Strategies for Comprehending Multiple Texts with Conflicting Information**

We assumed that three metacognitive strategies that strengthen the comprehension of belief-inconsistent text information are particularly important to counter the text-belief consistency effect (Richter & Schmid, 2010). First, if routine validation processes are one source of the text-belief consistency effect, learners' awareness of their prior beliefs and their biasing influence on comprehension become relevant. Increased metacognitive awareness of one's beliefs and their role in comprehension is also a precondition for deliberating on the plausibility of belief-inconsistent information. Hence, becoming aware of the biasing influence of one's prior beliefs was the first metacognitive strategy we considered relevant for the comprehension of multiple belief-consistent and belief-inconsistent texts.

In addition, following Baker's (1985) distinction between the internal and external consistency standards, two further metacognitive strategies were deemed as highly relevant for multiple text comprehension. The second metacognitive strategy consisted of identifying textual inconsistencies (internal consistency standard). When reading multiple documents, identifying inconsistencies should include both, detecting inconsistencies within and across documents. When processing multiple texts with conflicting information, the detection of intertextual inconsistencies can be fostered by actively judging the plausibility of the arguments in one text from the perspective of the other text.

In contrast, the use of the external standard is directed at relationships between prior knowledge and text information (Baker, 1985). In the context of reading belief-consistent and belief-inconsistent texts comprehension, monitoring violations of one's knowledge and beliefs appears to be crucial. Applying the external standard during multiple text comprehension can be fostered by actively using one's prior knowledge to critically evaluate the plausibility of the claims made by both belief-consistent and belief-inconsistent texts. This evaluation should enhance an elaboration of belief-inconsistent information and also its integration into learners' mental representation of the scientific issue.

These three metacognitive strategies—awareness of the influence of prior beliefs, monitoring for intertextual inconsistencies and the use of prior knowledge for argument evaluation—should help students to evaluate the plausibility of web-based scientific information (Mason et al., 2010) and should lead to a stronger consideration of belief-inconsistent information during multiple text comprehension. Nonetheless, a lack of relevant knowledge about necessary metacognitive strategies is not the only possible cause of failures to regulate comprehension of multiple texts with belief-inconsistent information. Such failures might also be due to the fact that learners are not sufficiently motivated to invest the cognitive effort necessary for applying the metacognitive strategies just described. The cognitive effort required should be particularly high if these strategies are not (yet) well practiced.

**Motivation and (Multiple) Text Comprehension**

Motivation is the main prerequisite for learners to activate and sustain cognitive and metacognitive resources during text comprehension (see, for example, the concept of strategic reading, Paris et al., 1983; see also Schunk & Zimmerman, 2009). But what motivates learners to invest resources in their comprehension process? In general, social cognitive theory predicts that motivation and performance in a new task is influenced by past performance and performance feedback (Bandura, 1997; Bandura & Locke, 2003; Phillips, Hollenbeck, & Ilgen, 1996). From this perspective, intra-individual goal regulation can be explained by the link between goal attainment in a previous task and the setting of goals for subsequent tasks. Goals determine the performance levels that learners are trying to attain (Krenn, Würth, & Hergovich, 2013). After successfully meeting or exceeding a goal, that is, a desired level of performance, learners adjust their goals upward. They set a higher goal compared to their previous performance which “creates new motivating discrepancies to be mastered” (Bandura, 1997, p.131). In other words, attaining a goal by successfully executing relevant actions in a given task will create a positive discrepancy between past performance and the aspired standard. This discrepancy, in turn, leads to positive expectations of success (Bandura, 1997; Donovan & Williams, 2003; Ilies & Judge, 2005; Krenn et al., 2013; Phillips et al., 1996; Tolli & Schmidt, 2008). In contrast, negative feedback, i.e., failing to master a previous task, has been shown to lead to downward or no adjustment of future goals, such as maintaining or lowering the difficulty level or effort in the subsequent task (Donovan & Williams, 2003; Ilies & Judge, 2005; Krenn et al., 2013; Phillips et al., 1996; Tolli & Schmidt, 2008).

A longitudinal field study investigating goal setting of college athletes supported this general prediction (Donovan & Williams, 2003). This study found that individuals set lower goals after failing to meet the desired performance (e.g., negative feedback) and higher goals after positive feedback (for similar results, see Ilies & Judge, 2005). In the same vein, research on self-regulated learning suggests that subsequent performance depends on learning effectiveness feedback. In his social cognitive theory of self-regulation, Zimmerman (2000) proposed self-regulation as a cyclical open-loop process in which feedback on previous performance is used to proactively adjust subsequent performance and goal pursuit. Thus, mastery experiences should enhance a learner's belief to be capable of a specific action, but negative feedback should decrease a learner's success expectation in a specific learning situation (Bandura, 1986; Zimmerman, 2000).

In sum, this research suggests that learners' motivation to engage in more challenging tasks, such as using new metacognitive strategies during multiple text comprehension, should depend on prior performance feedback. Success feedback in previous tasks is likely to enhance learners' motivation for further engagement (e.g., applying metacognitive strategies to a subsequent task), whereas failure feedback should lower learners' motivation to engage in new challenging tasks.

**The Present Study**

The present study followed two related goals. First, we attempted to replicate the text-belief consistency effect found by Maier and Richter (2013a) on situation model strength in the comprehension of multiple texts with conflicting information. Second and more importantly, we aimed at testing the hypothesis that providing learners with information on relevant metacognitive strategies combined with favorable motivational circumstances will strengthen the processing of belief-inconsistent information during comprehension and thus should eliminate the text-belief consistency effect.

To this end, we conducted an experiment in which participants read two texts, one arguing for the belief-consistent and one arguing for the belief-inconsistent argumentative stance on the controversial topic of whether vaccinations are harmful or beneficial. Based on schema-theoretical assumptions and the idea of epistemic validation, we expected that in the control group, which did not receive information about metacognitive strategies, the situation model for the text communicating belief-consistent information would be stronger compared to the situation model for the text communicating belief-inconsistent information (text-belief consistency effect).

To pursue the second goal, three groups of learners all received information on the three metacognitive strategies 1) becoming aware of the influence of prior beliefs, 2) monitoring for intertextual inconsistencies and 3) using prior knowledge for argument evaluation, all of which should be particularly relevant for processing belief-inconsistent information. To investigate the role of motivation, participants' motivation to use the information about the metacognitive strategies during reading the multiple texts was varied between the three metacognitive strategy groups. In particular, two of the three metacognitive strategy groups received feedback regarding their performance, which was either a failure or a success feedback on participants' performance in a practice trial. The third group received no feedback about their performance in a practice trial. We assumed that only learners receiving success feedback in the practice trial would set themselves higher goals. That is, they would feel confident in using the metacognitive strategies during the subsequent test trial. In contrast, failure feedback in a practice trial would lead to downward goal revision and would thus make the use of the metacognitive strategies and the construction of balanced situation models less likely. Moreover, we expected that learners receiving only information on the metacognitive strategies without feedback should also be less likely to use the metacognitive strategies. This expectation is based on the assumption that such participants are also less likely to set higher goals, which would be necessary to implement the metacognitive strategies. In sum, we predicted that learners in the metacognitive strategy groups—even when receiving the same information about metacognitive strategies—would differ in the strength of the texts' situation models. Learners receiving success feedback should achieve equally strong situation models for both the belief-consistent and the belief-inconsistent text, whereas learners receiving failure feedback and learners receiving no feedback similar to the control group should achieve a stronger situation model for the belief-consistent text.

We also investigated the joint impact of metacognitive strategies and motivational conditions on the memory for text (propositional text base, Kintsch, 1988). First, we assumed that as predicted by the schema-pointer-plus-tag model and similar to findings from Maier and Richter (2013a), belief-inconsistent information should receive priority (tagged) status in the memory representation of the text (Graesser, 1981; Schank & Abelson, 1977). Moreover, we assumed that the metacognitive strategies, which are directed at a strategic evaluation of the plausibility of belief-consistent and belief-inconsistent information, should only affect the construction of the mental representation formed in multiple text comprehension. Thus, memory for text will be unaffected.

In addition, we investigated whether learners' reading patterns differed as a function of text-belief consistency and experimental condition. Longer reading times often indicate that readers spend more cognitive resources on strategic knowledge-based activities such as drawing inferences (Magliano, Trabasso, & Graesser, 1999). Reading times are also prolonged when readers attempt to resolve inconsistencies in a text (Rapp & Mensink, 2011). These types of elaborative processing should occur in particular when learners receive information about relevant metacognitive strategies and are motivated to use these strategies. Moreover, we expected the three metacognitive strategies taught in the experiment are particularly relevant for processing belief-inconsistent texts. Combining these two assumptions, we expected longer reading times for the belief-inconsistent text in the experimental condition with positive feedback.

Given that learners' commitment to their beliefs is likely to impact the processing of belief-consistent and belief-inconsistent information (Pomerantz, Chaiken, & Tordesillas, 1995), the effect of the extremity and certainty of learners' prior beliefs as the two key indicators of learners' commitment to their beliefs were assessed and included in all analyses. In addition, we included learners' reading skills (operationalized as the efficiency of basic cognitive processes in reading, Richter & van Holt, 2005) in the reading time analysis to control for skill-based individual differences in reading times.

# Method

## Participants

Eighty-five university students (66 women and 19 men) participated in the study. Their average age was 23.5 years (*SD* = 2.9) and they were majoring in different disciplines within the social sciences (e.g., sociology and education). All participants were native German speakers and received a small monetary reward (8 Euros per hour) for participation.

## Text Material

**Text material for the test trial**. Two texts discussing the risks and benefits of vaccinations were used as experimental texts in the test trial. Vaccination was selected as a topic because for this topic, an independent sample of university students in a pilot study ($N = 55$) had indicated an overall strong agreement with one argumentative position (pro position, *vaccinations are necessary and beneficial*) and only weak agreement with the contrary argumentative position (contra position, *vaccinations are unnecessary and harmful*; for similar results in an independent survey with 10,000 German respondents, see Stiftung Warentest, 2012). In the pilot study, participants read short statements of two conflicting argumentative stances in different scientific controversies and then reported the extent to which they agreed or disagreed to the statement (ratings on a scale from 0 (*do not agree*) to 6 (*fully agree*)). A paired samples *t*-test revealed that learners favored the pro vaccination position ($M = 4.16$, $SD = 1.40$) over the contra vaccination position ($M = 1.85$, $SD = 1.48$, $t(54) = 6.13$, $p < .001$, $d = 0.83$). Based on these results, we constructed one text arguing that vaccinations are necessary and beneficial (*pro vaccination text*) and one text claiming that vaccinations are unnecessary and harmful (*contra vaccination text*). In all other respects we tried to keep the texts as parallel as possible to ensure their comparability. Both texts were constructed based on the information from freely accessible websites from reputable German magazines (e.g., Spiegel Online, http://www.spiegel.de; Deutsches Ärzteblatt, http://www.aerzteblatt.de) that are likely to be consulted by the general public for informal learning on scientific topics. Moreover, writing style, structure and length of all texts were held strictly parallel. A short statement of the text’s major claim appeared at the beginning of the texts. Afterwards, the four key arguments consisting of a claim followed by supporting evidence were presented, separated by subheadings. At the end, a short summary of the arguments and a conclusion reflecting the text’s major claim was presented (a sample text translated into English is available from the authors upon request). The average length of the text was 888 words and the mean readability score was 49.5 (moderate difficulty, determined with the German adaption of the Flesch`s Reading Ease Index, Amstad, 1978). To ensure the comparability of the text content we conducted a pilot-test with an independent sample of 28 university students. In this test, students rated the two texts as understandable, providing high-quality arguments and representing a clear stance toward the issue (see Table 1 for an overview of the text characteristics).

**Text material for the practice trial.** For the practice trial we used two texts debating the causes of climate change. Similar to the test trial text material, the two texts used in the practice trial took opposing positions in the scientific controversy. Thus, one text argued that global warming was caused by mankind (315 words) and the other text claimed that natural phenomena are the causes of global warming (321 words).

**Comprehension Measure**

Comprehension was measured on the levels of the situation model and memory for text with 24 test items per text with a recognition-verification task (modified after Schmalhofer & Glavanov, 1986). In this task, participants decided first whether or not three different types of test items represent information explicitly provided in one of the texts (*recognition question*). Participants were informed that for a *yes* response to the recognition question the test sentence does not have to be a verbatim copy of a sentence from one of the texts, but rather needs to correspond in content to a sentence from the texts. In addition, participants judged whether or not three different types of test items contained information that matched the situation described by the texts (*verification question*). The three test items were paraphrases of text information, inferences matching the text content and distracters (48 test items in total, eight items per item type and text). Table 2 presents examples for the three types of test items. Paraphrase items contained information that was explicitly provided by the text. Hence, paraphrase items were constructed by replacing the key content words of a text sentence with synonyms and changing the word order of the sentence. Responses to the paraphrase items were used to assess the memory for text. In contrast, inference items contained information that was not explicitly provided by the text. Instead, these items were constructed on the basis of possible inferences that could be drawn from the texts. These items also reflected the text's argumentative stance, i.e., they were controversial in their main predication. Moreover, participants needed to infer this information to build an adequate mental representation of the text content. Responses to these items assessed the strength of the situation model. Finally, distracter items communicated information that was neither an explicit content of the text nor a sensible inference from the text. In other words, the information communicated by distracter items was not part of text memory or the situation model. Rather, these items shared some superficial content aspect from the text by providing additional information about vaccination.

The measure for the memory for text was based on the proportions of *yes* responses to the paraphrase items and the inference items in the recognition question. For this measure, the probit-transformed proportions of *yes* responses to the inference items (false alarms) in the recognition task were subtracted from the probit-transformed proportions of *yes* responses to the paraphrase items (hits) in the recognition task. This procedure yields comprehension scores corrected for response tendencies (similar to the signal detection measure *d'*, see Schmalhofer & Glavanov, 1986, for details). In the same vein, the measure for situation model strength was based on the proportions of *yes* responses to inference items and distracter items in the verification task. Similar to the measure for the memory for text, the probit-transformed proportions of *yes* responses to the distracter items (false alarms) were subtracted from the probit-transformed proportions of *yes* responses to the inference items (hits).

**Reading Times**

Reading times were recorded for each paragraph, summed up for each text and standardized by the number of syllables to account for differences in text length.

**Feedback Manipulation**

Learners' motivation to use metacognitive strategies was varied by the valence of an evaluative feedback in the practice trial. The external feedback provided participants with information about the actual response accuracy (percentage of correct responses) to the situation model question in the practice trial. Moreover, participants were provided with a prearranged standard of comparison to capture the quality of their performance in reference to the performance of other students. The standard of comparison was varied in accordance to the planned valence of the evaluative feedback. To define this standard, the mean accuracy of responses in the verification task was estimated in a norming study ($N$ = 39). Given the high percentage of correct responses in the norming study ($M$ = .88, $SD$ = .07), we used more difficult distracter items in the experiment proper. Given the relatively easy distracters in the norming study, participants in the failure condition were told that on average students' accuracy in this task was 88%. In addition, they were told that in the lower performance quartile students answered less than 85% correctly and in the higher performance quartile more than 93% were answered correctly. Hence, participants that did not reach at least 88% correct responses in the practice trail received negative performance feedback. In contrast, in the success condition the standard of comparison was lowered compared to the average performance of students in the norming study. In the success condition, participants were told that on average students' accuracy in this task is 67%. This reference score was used because it was the minimal score achieved by participants in the norming study. Similarly, they were told that in the lower quartile students answered less than 60% correctly and in the higher quartile students answered more than 70% items accurately. Under the assumption that most participants will have more than 67% correct responses in the practice task, this standard of comparison was supposed to lead to positive performance feedback. In sum, this procedure was used to ensure that although participants received feedback on their actual performance, the outcome feedback and thus learners' motivation to use metacognitive strategies was varied independently of participants' actual performance.

**Metacognitive Strategies**

Learners in the metacognitive strategies groups received information about three metacognitive strategies that were expected to foster the integration of belief-inconsistent information into the mental representation of a scientific topic. The first metacognitive strategy (becoming aware of the influence of prior beliefs) was used to make participants aware of their prior beliefs and their influence on the processing of text information. The second metacognitive strategy (monitoring for intertextual relationships and inconsistencies) focused on the detection of intra- and intertextual inconsistencies and thinking about argumentative relationships between texts. Finally, the third metacognitive strategy (using prior knowledge for argument evaluation) addressed the use of prior knowledge to critically evaluate the plausibility of the claims made by belief-consistent and belief-inconsistent texts and the detection of prior knowledge violations. The wording of each instruction is provided in Table 3 (translated into English).

## Learner Characteristics

**Extremity of prior beliefs.** Participants' prior beliefs concerning vaccinations were assessed with 10 statements (response categories ranging from 1 (*totally disagree)* to 6 (*totally agree*)). Five statements claimed that vaccinations are necessary and beneficial (*pro vaccination belief scale*, e.g., "I think that vaccinations are the most important and most effective method against infectious diseases", Cronbach's α = .77) whereas the other five statements argued that vaccinations are unnecessary and harmful (*contra vaccination belief scale* e.g., "I am against vaccinations because they might overstrain my immune system", Cronbach's α = .79). The difference in the mean agreement to the two belief scales (mean agreement to pro vaccination belief scale – mean agreement to contra vaccination belief scale) served as indicator of the extremity of participants' prior beliefs.

**Certainty of prior beliefs**. We also assessed the certainty of participants' beliefs (response categories ranging from 1 = *very uncertain* to 6 = *totally certain*). To this end, participants stated for the ten belief-items how certain they were about their dis-/agreement to the belief item directly after responding to the corresponding belief item ("Please indicate how sure you are in your agreement to the former statement ", Cronbach's α = .83).

**Reading skills**. Reading skills were assessed with the sentence verification subtest of the German reading test ELVES (Cronbach's alpha = .87, Richter & van Holt, 2005). In this task, participants were asked to judge whether simple assertions about abstract and concrete concepts are true or false. Test scores combine the accuracy and the speed of a given response for each item (see Richter & van Holt, 2005, for details). Thus, this subtest measures learners' efficiency in using propositional strategies for reading comprehension on the sentence level (i.e., lexical access; syntactic and semantic integration). Reading skill was included as a covariate in the analysis of reading times.

**Procedure**

Approximately one week prior to the experiment proper, participants' prior beliefs (extremity and certainty) and their reading skills were assessed. The experiment consisted of a practice trial and a test trial that were each divided into a reading phase, a recognition task, and a verification task. During the reading phase in the practice trial, participants read two scientific texts on climate change in a self-paced fashion, one paragraph at a time on a computer screen. Participants were instructed to read the texts carefully for comprehension. After participants completed the reading phase, the corresponding test items were presented in the recognition-verification task of the practice trial. In the recognition task, participants indicated for each test item whether or not the sentence was explicitly provided in one of the texts. In the subsequent verification task, participants judged whether or not the test item matched the situation that was described in the texts. Participants responded by pressing one of two response keys (marked green for *yes* and red for *no*). The sentences were presented one at a time in black letters (font type Arial, average height 0.56 cm, bold) on a white background and in random order. Immediately following the recognition-verification task, participants in the two feedback groups received feedback regarding their performance in the task. One group received feedback that compared their actual performance to a high standard of comparison (failure feedback), whereas the other group received feedback that compared their performance to a lower standard of comparison (success feedback). Afterwards, the two feedback groups as well as the third metacognitive strategies group received information about the three metacognitive strategies and were told to monitor the strategic influences on multiple text comprehension during the test trial. Participants in the no training group neither received feedback nor information about the three metacognitive strategies. After the practice trial, all participants completed the test trial. Similar to the practice trial, the test trial consisted of a reading phase (assessment of reading times for each paragraph) and a recognition-verification task (assessment of memory for text and situation model). At the end of the experiment, participants were thanked and debriefed.

**Design**

The experimental design was a 2 (*text-belief consistency*: consistent vs. inconsistent; varied within participants) X 4 (*experimental condition*: control group vs. knowledge about metacognitive strategies without feedback vs. knowledge about metacognitive strategies with negative feedback vs. knowledge about metacognitive strategies with positive feedback; varied between participants) design. The text order (consistent-inconsistent vs. inconsistent-consistent) was counterbalanced between participants and included as a control factor. In all analyses, the extremity of participants' prior beliefs and the certainty of their prior beliefs were included as covariates. In addition, for the analyses of the reading times, reading skills were included as additional covariate.

## Results

The hypotheses pertaining to effects of text-belief consistency and experimental condition were tested with an ANCOVA for designs with between- and within-subjects factors. The reading order of the texts was included as a control factor in the analysis. Moreover, the extremity of participants' prior beliefs (difference in agreement to belief scales, *z*-standardized) as well as the certainty of prior beliefs (*z*-standardized) were included as covariates. In the ANCOVA for the reading times as dependent variable, learners' *z*-standardized scores in the reading skill measure (sentence verification test) were included as a covariate.

The application of ANCOVA with non-varying slopes between groups is based on the assumptions that the residuals are distributed normally and that the regression slopes of the covariates should be equal in the different groups (Maxwell & Delaney, 2000, Ch. 9). The first assumption was fully met in the present data set. The distributions of the residuals of all dependent variables did not differ significantly from a normal distribution (Kolomogorov-Smirnov tests: $Z < 0.89$, $p > .41$). The second assumption of the homogeneity of slopes was tested by including the interactions of the covariates with the experimental groups. With one exception, these analyses did not indicate heterogeneity of slopes (all $F < 1$), implying that the second assumption was met as well. The exception occurred in the ANCOVA for reading times as the dependent variable with the covariate difference in agreement to belief scales, which marginally failed to reach significance, $F(3,67) = 2.4$, $p = .08$. To be on the safe side, we included the interaction term of this covariate with experimental condition in the ANCOVA of the reading times, allowing the slopes to differ between the experimental groups (see Maxwell & Delaney, 2000, pp. 406-420).

All hypothesis tests were based on type I error probability of .05. Under the assumptions of a medium effect size ($f = .25$ according to Cohen, 1988) and medium correlations ($\rho = .5$) between the levels of the independent variables in the population, the design and sample size of the experiment yielded a power ($1-\beta$) of .98 for detecting the focal interaction of text-belief consistency and experimental condition (power was computed with the software G*Power 3; Faul, Erdfelder, Lang, & Buchner, 2007). Descriptive statistics and intercorrelations of all variables are provided in Table 4. For clarity of presentation, occasional significant results for control variables which are irrelevant for our hypotheses are not reported (the full set of results is available from the authors upon request).

**Manipulation Check for Text-Belief Consistency**

Before analyzing the effects of text-belief consistency and the experimental condition on situation model strength and the memory for text, we investigated participants' initial beliefs. A paired sample *t*-test revealed that participants overall more strongly agreed with the argumentative position that vaccinations are necessary and beneficial ($M = 4.00$, $SD = 1.03$) but tended to disagree with the position that vaccinations are unnecessary and harmful ($M = 2.53$, $SD = 1.04$), $t(84) = 7.12$, $p < .001$, $d = 0.77$. However, a closer investigation of the differences in agreement to the belief scales revealed that 17 participants more strongly agreed to the claim that vaccinations are unnecessary and harmful. For this reason, text-belief consistency was defined on an individual basis. Accordingly, we defined the text to whose argumentative position learners more strongly agreed to as indicated by the prior belief scales, as the belief-consistent text. For instance, for participants who believed that vaccinations are harmful, i.e., which more strongly agreed with the contra belief scale, the text arguing for this argumentative position was considered as belief-consistent text.

**Validation of Evaluative Feedback**

Participants in the feedback conditions received feedback for their performance accuracy for the situation model question during the practice trial compared to the average accuracy of a comparison group. To vary the feedback condition independently of learners' actual performance in the task, the standard of comparison was decreased (success) or increased (failure). However, learners could still have been better or worse than the provided standard of comparison. To investigate if the planned valence of the feedback (success vs. failure) was consistent with the actual valence, we examined whether participants' performance was lower or higher than the standard of comparison in the respective condition. In the success condition, the performance of six participants was lower than the average performance of the comparison group, yielding a failure feedback. In the failure condition, the performance of one participant was better than the average performance of the comparison group, yielding a success feedback. Given that the actual feedback (success vs. failure) did not result in the planned feedback condition for these participants, the seven outliers were reclassified.

**Results for Situation Model Strength**

In the analysis of situation model strength, the ANCOVA revealed a main effect of text-belief consistency, $F(1, 75) = 15.8$, $p < .001$, $\eta_p^2 = .17$. The situation model for the belief-consistent text was stronger overall ($M = 1.75$, $SE_M = 0.09$) than the situation model for the belief-inconsistent text ($M = 1.39$, $SE_M = 0.08$). However, the effect of text-belief consistency was moderated by experimental condition, $F(3, 75) = 2.7$, $p < .05$, $\eta_p^2 = .10$ (Figure 1). In line with our assumptions for the control group, the situation model for the belief-consistent text ($M = 1.59$, $SE_M = 0.15$) was stronger than the situation model for the belief-inconsistent text ($M = 1.26$, $SE_M = 0.14$), $F(1, 75) = 4.2$, $p < .05$, $\eta_p^2 = .05$. Similarly, in the metacognitive strategy group with failure feedback, learners' situation model for the text communicating belief-consistent information ($M = 1.76$, $SE_M = 0.15$) was stronger than the situation model for the text communicating belief-inconsistent information ($M = 1.37$, $SE_M = 0.14$), $F(1, 75) = 5.9$, $p < .05$, $\eta_p^2 = .07$. For the metacognitive strategy group without feedback, we found the same pattern, $F(1, 75) = 16.7$, $p < .01$, $\eta_p^2 = .18$. Participants in this group had a stronger situation model for the text communicating belief-consistent information ($M = 1.88$, $SE_M = 0.18$) than for the text communicating belief-inconsistent information ($M = 1.12$, $SE_M = 0.16$). In contrast, in the metacognitive strategy group with success feedback, there was no significant difference in the strength of the situation model for the belief-consistent text ($M = 1.74$, $SE_M = 0.20$) and the situation model for the belief-inconsistent text ($M = 1.80$, $SE_M = 0.19$), $F(1, 75) = 0.7$, $p = .78$. To sum up, in line with the hypotheses, we found a text-belief consistency effect in all experimental conditions except for the group that received information about metacognitive strategies and positive performance feedback. In the latter group, participants were able to construct a strong situation model for the belief-consistent as well as the belief-inconsistent text.

**Results for Memory for Text**

For the memory for text, the main effect of text-belief consistency reached significance, $F$ (1,75) = 4.0, $p$ = .05, $\eta_p^2$ = .05. Overall, participants' memory for the belief-inconsistent text was stronger ($M$ = 0.63, $SE_M$ = 0.08) than their memory for the belief-consistent text ($M$ = 0.44, $SE_M$ = 0.08). However, there was an interaction between text-belief consistency and experimental condition, $F$ (1,75) = 3.4, $p$ < .05, $\eta_p^2$ = .12 (Figure 2). In the control group, participants' memory for the belief-inconsistent text was slightly stronger ($M$ = 0.60, $SE_M$ = 0.14) than for the belief-consistent text ($M$ = 0.40, $SE_M$ = 0.15) but this difference failed to reach significance, $F$ (1,75) = 1.3, $p$ = .26. For the metacognitive strategy group without feedback, as well as the metacognitive strategy group with negative feedback, there was no difference in the memory for the belief-inconsistent (without feedback: $M$ = 0.64, $SE_M$ = 0.16; negative feedback: $M$ = 0.39, $SE_M$ = 0.14) and the belief-consistent text (without feedback: $M$ = 0.65, $SE_M$ = 0.17; negative feedback: $M$ = 0.55, $SE_M$ = 0.15), all $F$'s < 1.0, nonsignificant. In contrast, participants who received the metacognitive strategies with positive feedback had a stronger memory for the belief-inconsistent text ($M$ = 0.90, $SE_M$ = 0.18) than for the belief-consistent text ($M$ = 0.15, $SE_M$ = 0.20), $F$ (1,75) = 10.4, $p$ < .01, $\eta_p^2$ = .12.

**Results for Reading Times**

Reading times per syllable deviating more than two standard deviations from the mean of the experimental condition (as defined by our experimental design) were discarded from the analysis (4.2% of the data). The repeated-measurements ANCOVA on reading times revealed a main effect of text-belief consistency, $F$(1, 70) = 10.7, $p$ < .01, $\eta_p^2$ = .13 (Figure 3). In general, participants spent more time reading the belief-inconsistent text ($M$ = 178.4 ms/syllable, $SE_M$ = 4.9) than the belief-consistent text ($M$ = 172.4 ms/syllable, $SE_M$ = 4.7). Moreover, there was a main effect of training condition, $F(3, 70) = 3.3$, $p < .05$, $\eta_p^2 = .12$. The reading times increased monotonically from the control group ($M = 153.8$ ms/syllable, $SE_M = 8.1$) over the groups that received information on metacognitive strategies combined with no feedback ($M = 176.2$ ms/syllable, $SE_M = 9.6$) or failure feedback ($M = 177.5$ ms/syllable, $SE_M = 7.8$) to the group that received information about metacognitive strategies plus the positive feedback ($M = 194.1$ ms/syllable, $SE_M = 11.6$).

These main effects were qualified further by an interaction between text-belief consistency and the training condition, $F(3, 70) = 3.4$, $p < .05$, $\eta_p^2 = .13$. In line with the hypothesis, participants in the control group spent a comparable amount of time reading the belief-consistent ($M = 155.1$ ms/syllable, $SE_M = 8.1$) and the belief-inconsistent text ($M = 152.4$ ms/syllable, $SE_M = 8.4$), $F(1, 70) = 0.8$, $p = .38$. In contrast, participants who received information about metacognitive strategies and positive feedback spent longer reading the belief-inconsistent texts ($M = 198.4$ ms/syllable, $SE_M = 12.0$) than the belief-consistent texts ($M = 189.7$ ms/syllable, $SE_M = 11.6$), $F(1, 74) = 7.7$, $p < .05$, $\eta_p^2 = .05$. Unexpectedly, a similar difference occurred in the group that received negative feedback (belief-consistent texts: $M = 173.2$ ms/syllable, $SE_M = 7.8$; belief-inconsistent texts: $M = 181.8$ ms/syllable, $SE_M = 8.1$; $F(1, 70) = 8.3$, $p < .01$, $\eta_p^2 = .11$) and in the group that received no feedback (belief-consistent texts: $M = 171.7$ ms/syllable, $SE_M = 9.5$ ms; belief-inconsistent texts: $M = 180.8$ ms/syllable, $SE_M = 9.9$), $F(1, 70) = 6.2$, $p < .05$, $\eta_p^2 = .08$.

Thus, the hypothesis concerning reading times was only partially corroborated: Reading times were prolonged for the belief-inconsistent texts in all groups that received information on metacognitive strategies, regardless of the nature of performance feedback.

## Discussion

The primary purpose of the present study was to investigate the extent that information about metacognitive strategies and motivation to use these strategies moderate the text-belief consistency effect in multiple text comprehension. Results revealed that learners in the control group and learners who had received information about metacognitive strategies in combination with a failure feedback in a preliminary practice trial or no feedback built a stronger situation model for the text communicating belief-consistent information compared to the text with belief-inconsistent information (text-belief consistency effect). In contrast, learners receiving information about metacognitive strategies in combination with success feedback achieved equally strong situation model representations for the texts with belief-consistent and belief-inconsistent information. We also found that learners receiving information about the metacognitive strategies in combination with success feedback had a stronger memory for the belief-inconsistent text. In addition, analyses of the reading times revealed that learners in the control group spent equal amounts of time reading belief-consistent and belief-inconsistent texts, whereas learners in the groups that received information about the metacognitive strategies devoted significantly more time to read the belief-inconsistent text.

The observed text-belief consistency effect in the control group and in the metacognitive strategy group with no or failure feedback is a replication of earlier findings (Maier & Richter, 2013a). Apparently, prior beliefs bias the situation model in the comprehension of multiple texts and can lead to a one-sided mental representation of a controversial scientific issue. Two types of processes might underlie this effect. One possibility is that prior beliefs serve as schemata and facilitate the integration of belief-consistent information into long-term memory (Schank & Abelson, 1977; Smith & Graesser, 1981). Accordingly, belief-inconsistent information should be remembered better because it cannot be integrated, resulting in a prominent memory representation. Such a reverse text-belief consistency effect should have occurred in the control group (replicating the results found by Maier & Richter, 2013a), but it was not significant. However, a trend was found in the expected direction so that the failure to establish a reverse text-belief consistency effect might be due to low power associated with the relatively small size of the control group.

Another possibility—compatible with the schema-theoretic explanation—is that the advantage of belief-consistent information results from using prior beliefs to validate the plausibility of incoming text information as an integral part of text comprehension (epistemic validation; Richter et al., 2009; Singer, 2006). Epistemic validation might lead to perceiving the belief-inconsistent information as less plausible with the consequence that it is less likely to be integrated into learners' situation model (Schroeder et al., 2008; see also Lombardi et al., 2013).

Our results further suggest that learners can overcome the text-belief consistency effect by using metacognitive strategies that directly address the challenges of comprehending multiple belief-consistent and belief-inconsistent texts, provided that they are motivated to accept these challenges. Comprehension benefits gained from metacognitive processing are widely accepted (e.g., Zimmerman, 2000), especially in ill-structured learning domains such as multiple text comprehension (Stadtler & Bromme, 2007), and motivation is crucial for successfully applying these strategies (e.g. Zimmerman, 2000). The present study advances this research by testing the effectiveness of three metacognitive strategies that were designed specifically to assist learners in a rational evaluation of information to overcome myside biases at the comprehension stage. Remarkably, under favorable motivational conditions, providing concise information on the three metacognitive strategies was sufficient to completely eliminate the text-belief consistency effect. These conditions not only yielded a strong and balanced situation model but also a better memory for information explicitly provided by the belief-inconsistent texts. This suggests that the metacognitive strategies might cause readers to direct more attention to belief-inconsistent information, creating a better propositional text base that serves as the basis for elaborative processes. However, it is important to note that learners need to be motivated to engage in the use of metacognitive strategies for processing belief-inconsistent information. All three metacognitive strategies are likely resource-demanding, such that learners need to be willing to spend the necessary cognitive effort to implement the strategies. As a consequence, teaching oneself metacognitive strategies is not sufficient to reduce the text-belief consistency effect. Instead learners need to be given the expectation that they are likely to succeed in engaging in the new and challenging task to apply the strategies when reading. Performance feedback is not the only way to induce such a favorable motivational state. For example, teachers can use the positive-discrepancy assumption proposed by the social-cognitive theory (Bandura, 1997) in such a way that goal attainment leads to new and more ambitious goals which in turn motivates the engagement in new cognitively challenging tasks.

The reading time data partly supported the interpretation provided for the comprehension data. Reading times were longer for belief-inconsistent compared to belief-consistent texts for participants receiving information on the metacognition strategies and positive performance feedback. Unexpectedly, however, the same effect occurred for all participants receiving information about the metacognitive strategies, regardless of whether they received positive, negative, or no performance feedback. Thus, the information on metacognitive strategies seems to have increased learners' processing of belief-inconsistent information, but this increase only paid off in terms of a better situation model when the information about metacognitive strategies was paired with favorable motivational conditions. Given that reading times are a rather gross indicator of cognitive effort, the explanation for this pattern of effects remains unclear at this point. One possibility is that participants in the metacognition groups with failure feedback and no feedback made some attempts to apply the new strategies to reading the texts with conflicting information. However, these attempts might have been inconsistent because the participants in the groups with no feedback or failure feedback lacked the expectation that the metacognitive strategies would really increase their comprehension performance (similar to the *production deficit* observed in young children's use of memory strategies, Hasselhorn, 1996). This interpretation is backed up by the observation that reading times were longest overall in the metacognition group that received positive feedback. More informative indicators of cognitive processes such as think-aloud data might help to clarify this issue in future studies.

It is important to note that reducing the text-belief consistency effect is not an end in itself. When reading about currently debated scientific controversies for which there is no definitive right or wrong, acknowledging and comprehending arguments from both sides of the controversy is important. For example, the question whether or not vaccinations are harmful or beneficial is critically discussed among scholars and scientists, but this scientific issue is also highly relevant for everyday decision making of lay people (e.g., deciding whether or not one should get vaccinated). The rationality of such decisions is hampered by the text-belief consistency effect. Thus, achieving a balanced representation of controversially debated scientific topics is an important educational goal. Reducing the text-belief consistency effect by increasing learners' understanding of belief-inconsistent information should lead to a comprehensive understanding of the scientific issue and possibly also to a justifiable point of view. Note, though, that the latter aspect was not investigated directly in the present study. Furthermore, we cannot draw any conclusions about the extent learners resolved the inconsistencies and related the contradictory information to each other. This issue should also be investigated in future studies, for example, by using tasks that require inferences across texts (Strømsø et al., 2010).

Even though the present study provides new insight into the comprehension of multiple texts, several limitations need to be addressed. First, we were interested in the combined effectiveness of the three metacognitive strategies. The first strategy, becoming aware of prior beliefs and their influence on comprehension processes, appears to be the most fundamental of the three strategies, because insight in potential belief biases is a precondition for monitoring one’s comprehension process for the occurrence of such biases and for taking regulatory actions. The second strategy, checking for intertextual relationships and inconsistencies, encompasses comprehension monitoring according to the internal standard of consistency (Baker, 1985) but extends this notion to monitoring inconsistencies and argumentative relationships between multiple texts. The strategy overlaps with the strategy of corroboration, which Wineburg (1991) identified as one of the strategies expert historians use to extract factual information on historical events out of multiple documents (see also Britt & Aglinskas, 2002). Nonetheless, the intertextual relationships that can exist between multiple text documents on science-related topics clearly differ greatly from those relevant for history (e.g., the relationships of theoretical claim to evidence and counter-evidence are specific to science-related texts). The third strategy, using prior knowledge for a thorough evaluation of arguments, resembles the intentional application of Baker's (1985) external standard of consistency but also adds a new aspect. Rather than merely scrutinizing the texts to detect information that is inconsistent with prior knowledge, the strategy also implies an active use of prior knowledge to resolve the inconsistencies. This type of processing can be accomplished, for example, by seeking an explanation that accommodates the inconsistent propositions (Walsh & Johnson-Laird, 2009). However, given that learners received information on three different metacognitive strategies, we cannot be sure which strategy might have been especially helpful for learners to comprehend the belief-inconsistent text. We assume that a combination of all three metacognitive strategies is particularly effective but this issue has not been resolved at this point. Therefore, further research should investigate the role of the three metacognitive strategies to assess how they act in concert in the comprehension of multiple texts with conflicting information.

Another limitation is that the present study focused on performance feedback as a means to increase the motivation to use the metacognitive strategies during comprehension. The effect of such an intervention is likely to be short-lived. The impact of motivational factors that foster long-term learning with belief-consistent and belief-inconsistent multiple texts remains an open question that should be addressed in future research. Moreover, in our study learners were given only one belief-consistent and one belief-inconsistent text. One reason for using only two texts was that such a minimalistic design would allow a stronger test for the effectiveness of the metacognitive strategies in combination with positive feedback. Although the text-belief consistency effect has already been established in experiments with four experimental texts (Maier & Richter, 2013a), the moderating role of metacognitive strategies (plus motivation) is a new finding. For this reason, investigating the impact of text-belief consistency and metacognitive strategies in a learning scenario more typical of multiple text comprehension in which belief-consistent and belief-inconsistent information is spread across several texts and sources is likely to advance the literature on this topic.

In sum, the present study provides additional support for the assumption that prior beliefs bias the comprehension of multiple texts about controversial scientific issues. However, it also suggests a means by which a text-belief consistency effect in multiple text comprehension can be reduced through simple information about metacognitive strategies when confronting belief-inconsistent information. Such information must be accompanied by measures to ensure that learners are also sufficiently motivated to invest the cognitive effort needed to implement these strategies. Given that learners were only provided with information about the metacognitive strategies without direct training, the results suggest that strategies—if they are used—effectively boost the comprehension of belief-inconsistent texts. Assessing whether the effectiveness of the metacognitive strategies can be enhanced when they are directly trained would be informative. Comprehending multiple texts about science topics is a highly relevant but difficult task for many learners (including university students). Thus, future studies should focus on the training of multiple text comprehension when assessing the influence of learners' prior beliefs, their knowledge about metacognitive strategies and also their motivation.

**References**

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

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

Baker, L. (1985). Differences in the standards used by college students to evaluate their comprehension of expository prose. *Reading Research Quarterly, 20*, 297-313.

Baker, L. (1989). Metacognition, comprehension monitoring and the adult reader. *Educational Psychology Review, 1*, 3-38.

Baker, L., & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Reading Research Quarterly, 17*, 281-294.

Baker, L., & Beall, L. C. (2009). Metacognitive processes and reading comprehension. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 373-388). New York: Routledge.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353-394). New York: Longman.

Bandura, A. (1986). *Social Foundations of thought and action.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*, 87–99.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 209-233). Mawah, NJ: Erlbaum.

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction, 20*, 485–522.

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.

Donovan, J. J., & Williams, K. J. (2003). Missing the mark: Effects of time and causal attributions on goal revision in response to goal-performance discrepancies. *Journal of Applied Psychology, 88*, 379–390.

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

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-235). Hillsdale, NJ: Erlbaum.

Graesser, A.C. (1981). *Prose comprehension beyond the word*. New York: Springer.

Hasselhorn, M. (1996). *Kategoriales Organisieren bei Kindern: Zur Entwicklung einer Gedächtnisstrategie* [Categorial organization in children: On the development of a memory strategy]. Göttingen, Germany: Hogrefe.

Ilies, R., & Judge, T. A. (2005). Goal regulation across time: The effects of feedback and affect. *Journal of Applied Psychology, 90*, 453–467.

Isberner, M.-B., & Richter, T. (2013). Can readers ignore implausibility? Evidence for nonstrategic monitoring of event-based plausibility in language comprehension. *Acta Psychologica, 142*, 15-22.

Isberner, M.-B., & Richter, T. (in press). 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*. 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.

Kardash, C. A. M., & Scholes, R. J. (1995). Effects of pre-existing beliefs and repeated readings on belief change, comprehension, and recall of persuasive text. *Contemporary Educational Psychology, 20*, 201-221.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163-182.

Krenn, B., Würth, S., & Hergovich, A. (2013). The impact of feedback on goal setting and task performance. *Swiss Journal of Psychology, 72*, 79-89.

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

Lombardi, D. A. (2012). *Students' conceptions about climate change: Using critical* *evaluation to influence plausibility reappraisals and knowledge reconstruction.* Unpublished dissertation, University of Nevada, Las Vegas.

Lombardi, D., Sinatra, G. M., & Nussbaum, E. M. (2013). Plausibility reappraisals and shifts in middle-school students' climate change conceptions. *Learning and Instruction, 27*, 50-62.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91*, 615-629.

Mason, L., Boldrin, A., & Ariasi, N. (2010). Searching the web to learn about a controversial topic: Are students epistemically active? *Instructional Science, 38*, 607-633.

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

Maier, J. & Richter, T. (2013b). How nonexperts understand conflicting information on social science issues: The role of perceived plausibility and reading goals. *Journal of Media Psychology, 25*, 14-26.

Maxwell, S. E., & Delaney, H. D. (2000). *Designing experiments and analyzing data: A model comparison perspective*. Mawah, NJ: Erlbaum.

Myers, M., & Paris, S. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology, 70*, 680-690.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175–220.

Nokes, J. D., Dole, J. A., & Hacker, D. (2007). Teaching high school students to use heuristics while reading historical texts. *Journal of Educational Psychology, 99*, 492-504.

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*, 293-316.

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99-122). Mawah, NJ: Erlbaum.

Phillips, J. M., Hollenbeck, J. R., & Ilgen, D. R. (1996). Prevalence and prediction of positive discrepancy creation: Examining a discrepancy between two self-regulation theories. *Journal of Applied Psychology, 81,* 498–511.

Pomerantz, E. M., Chaiken, S., & Tordesillas, R. S. (1995). Attitude strength and resistance processes. *Journal of Personality and Social Psychology, 69,* 408–419

Pratkanis, A. R. (1989). The cognitive representation of attitudes. In R. Pratkanis, S. Breckler, & S. Greenwald (Eds.), *Attitude structure and function* (pp. 70-98). Hillsdale, NJ: Erlbaum.

Rapp, D., & Mensink, M. (2011). Focusing effects from online and offline reading tasks. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp.141-164). Charlotte, NC: Information Age Publishers.

Richter, T. (2011). Cognitive flexibility and epistemic validation in learning from multiple texts. In J. Elen, E. Stahl, R. Bromme, & G. Clarebout (Eds.), *Links between beliefs and cognitive flexibilit*y (pp. 125-140). Berlin: Springer.

Richter, T., & Schmid, S. (2010). Epistemological beliefs and epistemic strategies in self-regulated learning. *Metacognition and Learning, 5,* 47-65.

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.

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.

Rouet, J.-F. (2006). *The skills of document use: From text comprehension to web-based learning*. Mahwah, NJ: Erlbaum.

Schank, R. C., & Abelson, R. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.

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.

Smith, D. A., & Graesser, A. C. (1981). Memory for actions in scripted activities as a function of typicality, retention interval, and retrieval task. *Memory and Cognition, 9*, 550-559.

Stadtler, M., & Bromme, R. (2004). Laypersons searching the WWW for medical information: The role of metacognition. In K. Forbus, D. Gentner, & T. Reiger (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (p. 1638). Mahwah, NJ: Erlbaum.

Stadtler, M., & Bromme, R. (2007). Dealing with multiple documents on the WWW: The role of metacognition in the formation of documents models. *Computer-Supported Collaborative Learning, 2*, 191-210.

Stiftung Warentest (2012). *Online-Umfrage Impfen: Ergebnisse im Detail* [Online survey on vaccinations: Detailed results] [Online document]. Retrieved June 20, 2012, from http://www.test.de/Online-Umfrage-Impfen-Ergebnisse-im-Detail-4371826-0/

Strømsø, H., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning & Instruction, 20,* 192-204.

Tolli, A. P., & Schmidt, A. M. (2008). The role of feedback, causal attributions, and self-efficacy in goal revision. *Journal of Applied Psychology, 93,* 692-701.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4,* 45-69.

Walsh, C. R., & Johnson-Laird, P. N. (2009). Changing your mind. *Memory & Cognition, 37,* 624-631.

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, *83*, 73–87.

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.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego, CA: Academic Press.

Table 1:

*Text characteristics of the experimental texts*

| | Length[a] | Readability[b] | Plausibility[c] | | Understandability[c] | | Number of arguments[c] | | Clarity of stance[c] | | Interestingness[c] | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Belief-inconsistent text | 921 | 50 | 4.10 | 0.92 | 4.27 | 0.78 | 3.85 | 0.88 | 4.86 | 1.18 | 4.29 | 1.21 |
| Belief-consistent text | 854 | 49 | 4.40 | 0.89 | 4.59 | 0.67 | 3.54 | 1.10 | 5.46 | 1.07 | 4.23 | 0.99 |

*Note*. Plausibility = measured with five items (response categories ranging from 0 (*not at all)* to 6 (*totally);* Cronbach's α = .79). Understandability = measured with nine items (response categories ranging from 0 (*not at all)* to 6 (*totally);* Cronbach's α = .80). Number of Arguments = number of identified arguments in an open answer question. Clarity of Stance and Interestingness were assessed with one item each (response categories ranging from 0 (*not at all)* to 6 (*totally*)).

[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 28 university students.

Table 2:

*Examples of test items used in the recognition-verification task (translated into English)*

| Sentence type | |
|---|---|
| Original text part | Additionally, a high vaccination rate offers the benefit of double protection. First, vaccinations provide direct protection against an epidemic infection. Second, herd immunity also protects those who could not be vaccinated such as infants or immune-compromised people, as emphasized by Prof. Battegay, chief physician of the department of Infectiology at the University of Basel. Herd immunity is defined as the protection of the whole population through the presence of vaccinated individuals. *In an unvaccinated population, the epidemic has a snowball effect*. However, with a large number of vaccinated people, the chance of avoiding wide-spread epidemic outbreaks is negligible. This protection is possible because vaccinated people do not excrete causative organisms, which in essence protects non vaccinated people and in turn prevents the distribution of epidemics. |
| Paraphrase | Infectious diseases spread in form of a chain reaction in an unvaccinated population. |
| Inference | People who choose not to be vaccinated endanger the health of elderly people and infants. |
| Distracter | In the case of a congenital immune deficiency, the decision for or against a vaccination with a live vaccine must be made individually. |

*Note*. In the original text part, the sentence used for the paraphrase item is italicized.

Table 3:

*Wording of the instructions on the three metacognitive strategies (translated into English)*

| Metacognitive strategy | Instruction |
|---|---|
| Becoming aware of the influence of prior beliefs | Become aware of your own beliefs regarding the scientific topic before you start reading the texts. Our prior beliefs often automatically determine whether we perceive incoming information as plausible or implausible. Thereby, information that is inconsistent with our prior beliefs may be rejected without further reasoning about the plausibility of this information. If you are aware of your prior beliefs, you are able to critically reason about the plausibility of the information rather than just rejecting it. |
| Monitoring for intertextual relationships and inconsistencies | Try to integrate the arguments from different texts into one overall picture of the scientific issue. Even if the texts argue for different standpoints in the controversy, their arguments might be related to each other with regards to content. Figure out which parts of the texts are in conflict with each other and which parts are consistent with each other. Become aware of the relationships between the texts and think about ways to combine the arguments with each other. Critically evaluate the arguments from one text in the light of the arguments of the other text. |
| Using prior knowledge for argument evaluation | Examine the quality of arguments in the text based on your prior knowledge. Prior knowledge is a good source to deliberately decide whether or not new information is plausible and credible. However, it is important to note that prior knowledge differs from mere opinion! Use the facts and evidence you know about the topic to critically evaluate the texts. Analyze all new information —even the information that might at first sight be in line with what you believe to be true. |

Table 4:

*Descriptive statistics and intercorrelations of independent variables, covariates, and dependent variables*

| | | | Correlations | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 Experimental condition | 0.44 | 0.91 | 1 | | | | | | | | | | | |
| 2 Metacognition I | 0.06 | 0.71 | .03 | 1 | | | | | | | | | | |
| 3 Metacognition II | -0.29 | 0.79 | -.23* | -.04 | 1 | | | | | | | | | |
| 4 Text order | 0.45 | 0.50 | -.01 | .09 | -.17 | 1 | | | | | | | | |
| 5 Extremity of prior beliefs | 1.47 | 1.89 | -.03 | .08 | .17 | .04 | 1 | | | | | | | |
| 6 Certainty of prior beliefs | 4.67 | 0.86 | -.06 | -.07 | .04 | .13 | .33** | 1 | | | | | | |
| 7 Reading skills | 18.83 | 5.53 | .01 | .19 | .21 | .02 | .10 | -.15 | 1 | | | | | |
| 8 Situation model strength (BC) | 1.76 | 0.75 | -.00 | -.07 | .04 | -.11 | -.04 | -.19 | .04 | 1 | | | | |
| 9 Situation model strength (BIC) | 1.37 | 0.74 | .01 | .06 | .24* | -.11 | -.29** | -.10 | .13 | .37** | 1 | | | |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 Memory for text (BC) | 0.47 | 0.74 | .08 | -.06 | -.14 | -.01 | -.04 | -.01 | -.14 | .10 | .05 | 1 | | | |
| 11 Memory for text (BIC) | 0.62 | 0.70 | .02 | -.08 | .22* | .17 | .15 | .07 | .18 | .20 | -.00 | .19 | 1 | | |
| 12 Reading times (BC) | 167.71 | 42.65 | .21 | -.04 | -.105 | -.06 | -.16 | -.15 | -.44** | .15 | -.03 | .20 | .19 | 1 | |
| 13 Reading times (BIC) | 172.58 | 46.43 | .29** | -.03 | -.14 | .13 | .04 | -.03 | -.44** | .07 | -.13 | .25* | .28* | .91** | |

*Note*. BC = belief-consistent text, BIC = belief-inconsistent text. Experimental condition: contrast coded, -1 = control group vs. 1 = experimental groups. Metacognition I: contrast coded, -1= metacognitive Strategies without feedback vs. 1= metacognitive Strategies with failure feedback. Metacognition II: contrast coded, -1= metacognitive Strategies without and with failure feedback vs. Metacognitive Strategies with positive feedback. Text order: contrast coded, -1 = belief-consistent/ inconsistent vs. 1 = belief-inconsistent /consistent). Situation model strength: biased-corrected proportion of correctly inferred inference items. Memory for text: biased-corrected proportion of correctly recognized paraphrase items. Reading skills: assessed with the sentence verification subtest of the ELVES (Richter & van Holt, 2005). Reading times: Reading times per syllable in ms.

* $p < .05$, ** $p < .01$ (two-tailed).

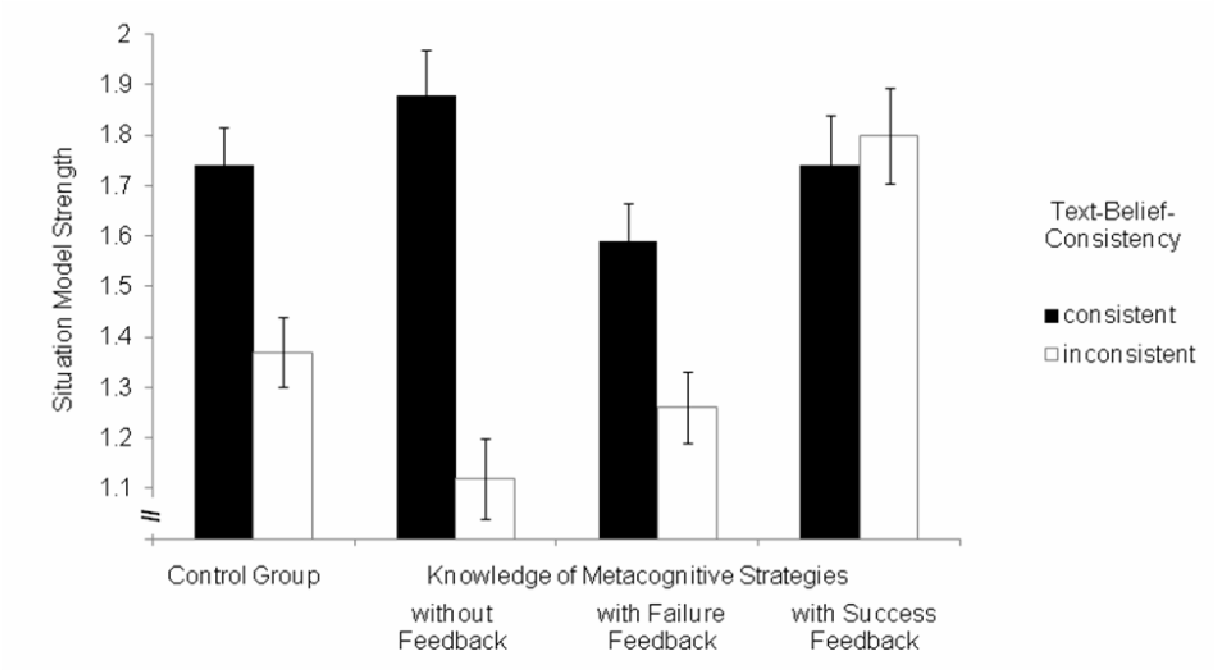


**Fig. 1** Interaction of text-belief consistency and experimental condition for the situation model (error bars represent the standard error of the mean).

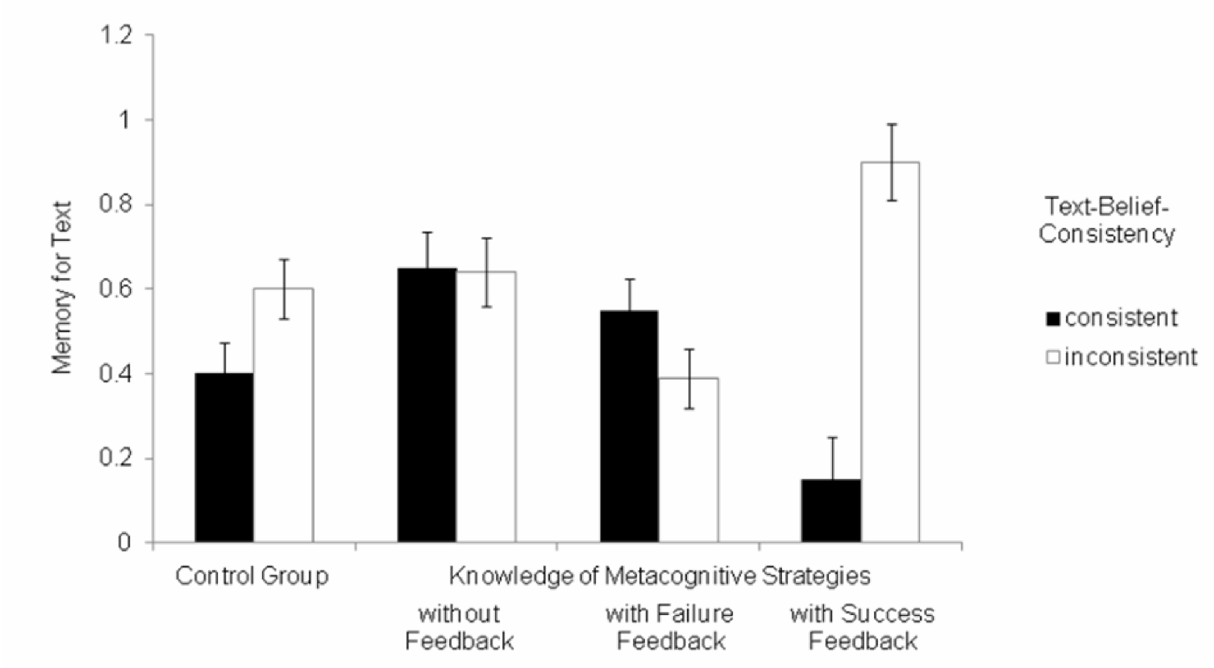


**Fig. 2** Interaction of text-belief consistency and experimental condition for the memory for text (error bars represent the standard error of the mean).

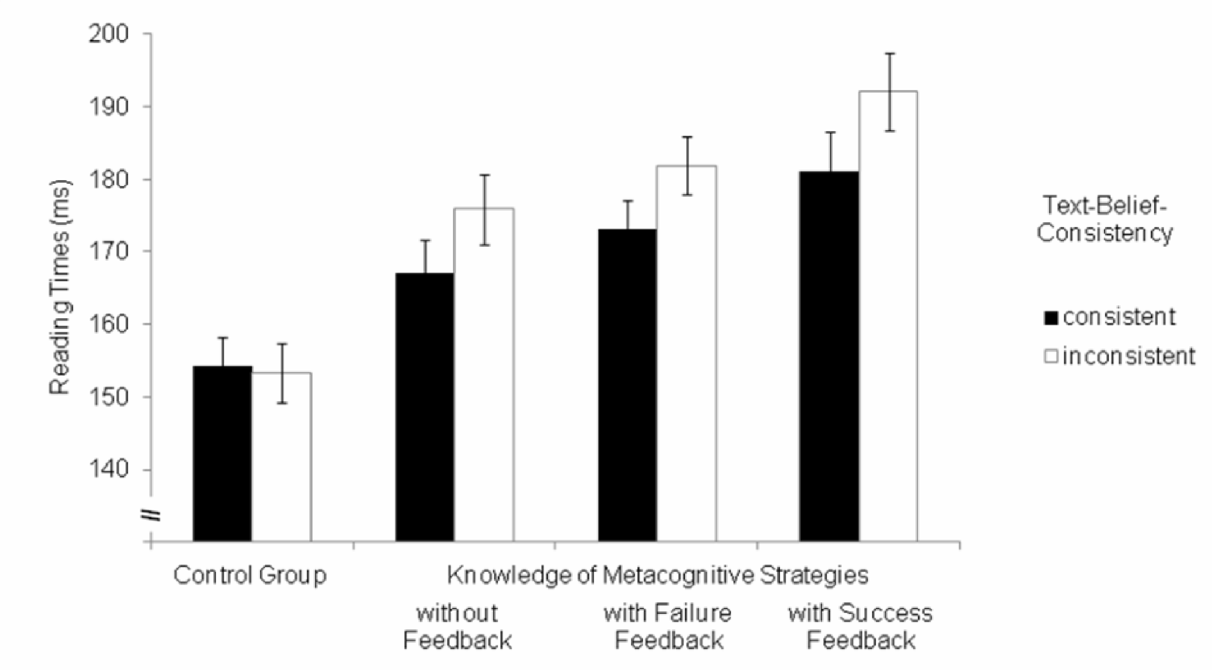


**Fig.3** Interaction of text-belief consistency and experimental condition for reading times (per syllable) in milliseconds (error bars represent the standard error of the mean).