How to understand what you don't believe: Metacognitive training prevents belief-biases in multiple text comprehension

Johanna Abendroth and Tobias Richter

University of Wuerzburg

Accepted for publication in the journal *Learning and Instruction* (2020)

Corresponding author:

Dr. Johanna Abendroth

University of Würzburg

Department of Psychology IV, Educational Psychology

Röntgenring 10

97070 Würzburg, Germany

E-mail: johanna.abendroth@uni-wuerzburg.de

Abstract

Prior beliefs often bias the comprehension of multiple science-related texts such that belief-consistent texts are better comprehended compared to belief-inconsistent texts (text-belief consistency effect). Two experiments were conducted to investigate whether a metacognitive strategy training focusing on belief-biasing validation processes can reduce the text-belief consistency effect in multiple text comprehension. Participants in the control condition received the PQ4R training as a well-situated and effective reading skill training that increases receptive elaborative processing. In Experiment 1 (n = 39) and Experiment 2 (n = 53) participants receiving a metacognitive strategy training achieved a similar level of comprehension of belief-consistent and belief-inconsistent texts, whereas a text-belief consistency effect was found in the PQ4R condition. These results indicate that a training focusing on belief-biasing validation processes prepare readers for the challenges to comprehend belief-relevant multiple texts, whereas strategies that foster receptive processing of information are not sufficient in the context of controversially discussed topics.

Keywords: validation, prior beliefs, metacognition, training

1. Introduction

The Word Wide Web has changed the way in which readers search, select, and comprehend science-related information. When reading online material, readers are frequently confronted with the challenging tasks to select, comprehend, and use a multitude of complex and often inconsistent and contrary texts on the same topic. Scientists use the Word Wide Web as a platform to discuss new theories and empirical results as quickly as they are developed, which is important for academic transparency and progress, but creates a challenging information landscape for laypersons. Such an open discussion of complex scientific debates can be currently observed on the topic of the COVID-19 global pandemic. Inconsistent information on the origin, diagnosis, treatment and most important prevention of the disease have been spread on the World Wide Web leading to a diversity of information on this issue of varying origins. Even if false information or intentional disinformation spread about the COVID-19 global pandemic is disregarded, a colorful portfolio of constantly changing facts and contrary arguments remains, which have a direct influence on the health of the individual and the global society. Understanding and comprehending such a multitude of information and multiple texts requires knowledgeable readers to critically evaluate scientific content to achieve their goals (Britt, Richter, & Rouet, 2014). Research suggests that prior beliefs play an important role during selecting, comprehending, and generating information from (multiple) texts. For example, McCrudden and Barnes (2016) found a that male secondary students provided more favorable evaluations of belief-consistent arguments. In addition to this belief effect on evaluation, readers have a strong tendency to select information that is in line with what they believe to be true (confirmation bias, Nickerson, 1998), comprehend textual information that is consistent with their prior beliefs to a greater extent than belief-inconsistent information (text-belief consistency effect; e.g., Eagly & Chaiken, 1993; Maier & Richter, 2013), maintain their beliefs even in the face of new information that explicitly corrects or discredits their beliefs (continued misinformation effect; e.g., Chinn & Brewer, 1993; Johnson & Seifert, 1994; Limon & Mason, 2002; Ross, Lepper, & Hubbard, 1975), and produce more belief-consistent compared to belief-inconsistent information when asked to generate arguments (myside bias; e.g. Wolfe, Britt, & Butler, 2009). These belief effects have in common a hampering effect on readers' ability to come to a full and comprehensive understanding of controversially discussed topics.

Given the strong influence of prior beliefs on the comprehension of (multiple) texts, the present study aimed at evaluating a metacognitive strategy intervention intended to assist readers in the comprehension of multiple science-related texts and to reduce belief biases in comprehension. The development of a metacognitive strategy intervention was informed by recent research indicating that text comprehension involves not only activation and integration of textual information but also routine epistemic monitoring of text-information with pertinent and accessible knowledge and beliefs (validation, Richter, 2015; Singer, 2013). Epistemic monitoring or validation (both terms are used interchangeably in this article) refer to a regular component of comprehension during which readers automatically check the consistency of new information with currently activated knowledge and beliefs as well as currently activated information from previously read text. Such nonstrategic validation processes are one likely source of belief effects in comprehension. We conducted two experiments to test the effectiveness of a metacognitive strategy training, focusing on belief effects and validation in multiple text comprehension. In Experiment 1, the method consisted of only a minimal training intervention, which was extended and developed into a computer-based training in Experiment 2. In both experiments, a learning scenario was created that is typical for learning with multiple documents. In detail, participants read two belief-relevant texts on the same issue that took rather divergent perspectives and provided conflicting arguments on the problem. The comprehension measure targeted the individual texts and not intertextual comprehension or integration between the two texts because the text-belief consistency effect refers to the relative strength of the situation models for individual texts that take divergent perspectives and provide conflicting information (e.g., Maier & Richter, 2013).

In the next paragraphs, we first describe the processes involved in multiple text comprehension and the Two-Step Model of Validation (Richter & Maier, 2017), which was formulated to explain belief-biases in the comprehension of multiple texts. We then discuss the instructions and interventions used to foster multiple text comprehension to provide the context for the training intervention examined in the present experiments.

1.1 Cognitive Processes Involved in Multiple Text Comprehension

Multiple texts comprehension is based on a variety of cognitive processes and results in different comprehension outcomes, from understanding the meaning of single texts to constructing the interrelations between texts. The MD-TRACE model (Multiple-Document Task-based Relevance Assessment and Content Extraction; Rouet & Britt, 2011) provides a depiction of the processing steps that readers need to undertake during multiple text comprehension. First, readers need to construct a task model based on their individual interpretations of the reading task, such as writing an argument or a summary. As a recent extension of the MD-TRACE model, in the RESOLV model (REading as problem SOLVing), readers not only construct a task model but also a context model, which is a representation of the social and physical environment (Rouet, Britt, & Durik, 2017). Both the MD-TRACE model and the RESOLV model assume that readers' internal resources, such as their prior knowledge and beliefs, influence the task model. In addition, the task model contains readers' schematic plans on how to achieve the subjective reading goal based on a benefit–cost analysis. In this process, the evolving task model provides a main structure for regulatory processes, in which reading activities are chosen and monitored with regard to the achievement of the specific reading goals. Thus, the task model is also used to select information based on its perceived relevance. That is, information perceived as relevant to achieve the established reading goal is selected more often.

Similar assumptions and observations have been made in research on perspective and relevance effects (e.g., Anderson & Pichert, 1978; McCrudden, Magliano, & Schraw, 2011) and on belief effects during multiple text comprehension (Maier, Richter, & Britt, 2018). For example, the goal-focusing model (McCrudden et al., 2011) proposes that personal and given intentions together form reading goals, which then influence processing and learning. Hence, similar to the assumption made in the MD-TRACE model, not only external cues (e.g., reading tasks) influence the reading goals set by readers and subsequently the attention and processing of goal-relevant information, but internal factors (e.g., beliefs) can also influence readers' task model. The interactive relationship between given intentions and personal intentions that together form reading goals and influence processing and learning is not unique for multiple text comprehension, but is relevant for text comprehension in general (McCrudden et al., 2011).

The MD-TRACE model further assumes that readers update their task product sequentially during reading. After reading, readers evaluate their outcome, that is, whether their understanding of the issue meets the task requirements they had established before reading. If readers are satisfied with the reading outcome, they will stop reading. Otherwise, they may continue to search for additional information. Similar to the task model, the evaluation of the reading outcome is assumed to be influenced by readers' internal resources such as their prior beliefs. A similar function of beliefs as epistemic gatekeeper has been discussed in the Two-Step Model of Validation (e.g., Richter & Maier, 2017), which will be discussed later. Moreover, the idea that readers evaluate their reading outcome is also related to the standard of coherence concept that determines whether readers are satisfied with their reading outcome or continue with strategic comprehension (van den Broek, Beker, & Oudega, 2015). In detail, readers' standard of coherence determines when readers are satisfied with their understanding in a particular reading situation. If this standard for sufficient comprehension is not met, readers then engage in additional strategic processing to meet the standard of coherence as an internally-generated standard.

As one product of reading multiple texts, readers need to build an intertext model. This intertext model comprises the semantic and argumentative relationships among the documents and relevant source information (Britt, Perfetti, Sandak, & Rouet, 1999; Perfetti, Rouet, & Britt, 1999; Rouet, Britt, Mason, & Perfetti, 1996). In addition, multiple text comprehension requires the construction of an integrated mental model. The integrated mental model represents the content information from multiple texts integrated with background knowledge and can vary on a continuum from unintegrated representations for each text to a fully integrated representation of all texts. For the construction of such an integrated mental model, it is essential that readers construct a situation model for each text content, that is, a more (or less) complete mental representation of what the text is about (Kintsch, 1988). The construction of the situation model requires readers to not only use information directly provided by the texts but also to elaborate on the information from information stored in their long-term memory. The construction of situation models largely relies on passive (automatic) cognitive processes. Textual information processed at a given point during reading passively activates the memory of information provided in the same text or in a different text read earlier and background knowledge and beliefs (e.g., Albrecht & O’Brien, 1993; Beker, Jolles, Lorch, & van den Broek, 2016; Perfetti, et al., 1999). To form the situation model, new textual information is integrated with the activated information from memory. The integration process is also a largely passive, text-driven process based on semantic associations between information from the text and information in memory (for an overview, see McNamara & Magliano, 2009).

Current research and theories also suggest that activation and integration during situation model construction are complemented by a third passive comprehension process, which has been termed validation (O’Brien & Cook, 2015; Richter, 2015; Richter & Singer, 2017). Validation is a routine cognitive process during which readers check the consistency of new information with currently activated knowledge and beliefs and currently activated information from previously read text(s). The idea that readers routinely and involuntarily validate information is supported by a number of studies based on different experimental paradigms, such as reading and reaction time and eye-tracking studies (for an overview, see Isberner & Richter, 2014). For example, Richter, Schroeder, and Wöhrmann (2009) presented single words on a computer screen, one-by-one in rapid succession (e.g., 300 ms per word), which together formed simple statements. When enough words were presented from a sentence to evaluate the veridicality of the sentence in the experimental trials, university students were asked to judge whether the last presented word was spelled correctly by providing a binary response (i.e., yes vs. no). Hence, the task was unrelated to the content of the sentence or even to the semantics of the word that required a response. Richter and colleagues (2009) investigated the response times (i.e., the time needed by participants to indicate whether the word was spelled correctly) and also error rates. The sentences were varied in such a way that the sentences were either true (e.g., libraries have books) or false (e.g., computers have emotions) based on easily accessible world knowledge. The results revealed that the validity of the sentence (was it true or false based on participants world knowledge) interacted with the required response of whether the last word was spelled correctly. When the sentence was false, but the correct response was positive (i.e.., the word was written correctly), participants needed more time to give their response, and they also made more errors (Richter, et al., 2009). These findings suggest that readers routinely and involuntarily validate information and that inconsistent, implausible, or false information evokes a negative response tendency, that is, a tendency to give a "no" response in an unrelated task. This effect has also been found for belief-relevant information (Gilead, Sela, & Maril, 2018). Gilead and colleagues (2018) found in an experiment with university students that belief-inconsistent claims (e.g., The internet has made people more isolated/sociable) elicited the same negative response tendency that was found for false or implausible statements. In the next section, we discuss possible effects of non-strategic validation processes on the comprehension of complex and conflicting multiple texts and the role of prior beliefs in this context.

1.2 Effects of Validation on Multiple Text Comprehension

Validation can be based on all information stored in long-term memory that becomes activated during comprehension (Richter, 2015). Prior beliefs are often strong enough to be passively activated during comprehension, especially when readers are reading science-related texts on topics that are relevant for everyday decisions such as the risks and benefits of vaccinations, electromagnetic radiation of cell phones, or nutritional side-effects. In such circumstances, readers often possess prior beliefs that are closer to one argumentative position in the science-related controversy, and these beliefs are as quickly used to evaluate argumentative claims as the information is comprehended (Maier et al., 2018; Voss, Fincher-Kiefer, Wiley, & Silfies, 1993; Wyer & Radvansky, 1999). For example, Voss and colleagues (1993) presented a combination of a claim (e.g., abortion should be illegal) and a reason (e.g., terminating a pregnancy is murder) to undergraduates. Claims were always presented first and reasons were either strong or weak and supporting or opposing. Participants' task was to judge whether the reason was supporting or opposing the presented claim. An important finding in this study was that participants were faster in responding to belief-consistent reasons, suggesting that a claim activates prior beliefs quickly during reading even when it was not required by the actual task.

The Two-Step Model of Validation (Richter & Maier, 2017) uses the concept of validation to explain how and why prior beliefs affect the comprehension of belief-relevant multiple texts. According to this model, readers by default often rely solely on the first of two validation steps, which is a routine non-strategic validation during comprehension (Step 1). In Step 1, readers' prior beliefs serve as an epistemic gatekeeper. This mechanism leads to an immediate disruption of comprehension when belief-inconsistent information is encountered because readers validate the consistency of textual information with their prior beliefs (Maier et al., 2018; Voss et al., 1993), but the disruption is typically not followed by readers' attempt to repair the inconsistency with strategic and resource-intensive comprehension. Consequentially, belief-inconsistent information is processed in a shallower manner than belief-consistent information. This bias in processing leads to poorer comprehension of belief-inconsistent information. In line with these assumptions, research with different groups of readers, different topics, and different comprehension tasks have repeatedly found that learning, memory, and retention of belief-relevant material is affected by prior beliefs. These studies specifically show a text-belief consistency effect in which readers demonstrate a stronger comprehension of belief-consistent texts and arguments (e.g., Abendroth & Richter, 2020a; Maier & Richter, 2013, 2014; Wiley, 2005, for a systematic overview see Richter & Maier, 2017). For example, Wiley (2005) found a stronger situation model for belief-consistent arguments when undergraduate students read multiple arguments that were either arguing for the belief-consistent or the belief-inconsistent position in a publicly debated controversy. This effect occurred only when arguments of the same type (i.e., belief-consistent or belief-inconsistent) were presented together but not when belief-consistent and belief-inconsistent arguments were presented interleaved.

The idea that beliefs serve as a type of filter to allocate cognitive resources to belief-relevant material during reading is similar to the idea of relevance (e.g., McCrudden, et al., 2011) or perspective effects (e.g., Anderson & Pichert, 1978). Research on these topics has repeatedly found that information perceived as relevant for a given perspective or reading goal receives more attention during reading and is better remembered after reading (e.g., Anmarkrud, McCrudden, Bråten, & Strømsø, 2013; Kaakinen & Hyönä, 2011). A similar effect might occur based on prior beliefs, that is, prior beliefs serve as standards of relevance affecting immediate and delayed processing as well as memory and comprehension of belief-relevant material. However, in contrast to relevance instructions as externally-provided cues, the influence of prior beliefs on processing and comprehension might be beyond readers' strategic control.

The comprehension of conflicting multiple texts requires that readers combine information from all texts into a complete picture of the issue to achieve a comprehensive understanding. Hence, this requirement should also include comprehending and understanding belief-inconsistent arguments and texts. The Two-Step Model of Validation proposes that readers will be successful in comprehending belief-inconsistent texts only when they engage in a second, optional validation step that includes strategic, resource-intensive, and goal-dependent elaboration of new information (Step 2). Processes involved in the strategic elaboration of information require a greater amount of cognitive resources, prior knowledge, and metacognitive knowledge to shield against belief effects. Moreover, such strategic validation processes should be especially beneficial when reading belief-inconsistent information because this type of information receives less cognitive processing by default. In support of this assumption, eye-tracking data (Maier et al., 2018) indicates that strategic processing and comprehension are linked. In detail, university undergraduates comprehended belief-consistent and belief-inconsistent multiple texts equally well when they performed more look-backs to earlier parts of the texts as strategic processing indicators. Time-pressure, low working memory capacity, low prior knowledge, or unawareness of belief effects make it unlikely that readers are able to successfully comprehend belief-consistent and belief-inconsistent multiple texts (Richter & Maier, 2017). Prior knowledge, for example, is one important pre-requisite for elaborative processing in the framework of the Two-Step Model of Validation. In contrast to prior beliefs, prior knowledge can be used to resolve inconsistencies between multiple texts and to successfully comprehend belief-inconsistent information. In line with this assumption, Wiley (2005) found that prior knowledge can shield against the text-belief consistency effect. In this study, undergraduates with low domain knowledge on a given topic (e.g., political science graduate students reading about a legal topic) were influenced by their beliefs in terms of a text-belief consistency effect in argument recall, whereas students with high domain knowledge (e.g., law students reading legal texts) showed no effect of prior beliefs on recall.

1.3 Interventions to Foster Multiple Text Comprehension

Multiple text comprehension is a challenging task for most readers because they usually experience difficulties in integrating information across texts (Bråten, Braasch, & Salmerón, in press; Richter & Maier, 2017; Stadtler, Bromme, & Rouet, 2018). Research has investigated several types of interventions that aim at increasing readers' comprehension of conflicting multiple texts (for overviews, see Barzilai, Zohar, & Mo-Hagani, 2018; Wiley, Jaeger, & Griffin, 2018). In research on multiple texts for which readers hold beliefs, such interventions have focused on specific reading goals (e.g., Bohn-Gettler & McCrudden, 2018; Maier & Richter, 2016; Wiley & Voss, 1999), the order in which texts were presented (Abendroth & Richter, 2020a; Maier & Richter, 2013; Wiley, 2005), or on metacognitive strategies (Maier & Richter, 2014). For example, Bohn-Gettler and McCrudden (2018) provided undergraduates with the instruction to either focus on pro or contra arguments of a dual-position text. This task-relevance instruction influenced memory such that participants recalled task-relevant information better, and this relevance effect in the comprehension outcome was independent of participants' beliefs. In contrast, strategic processing of the texts was influenced by participants' beliefs independent of the task instruction. In detail, participants used confirmation strategies when reading belief-consistent text and disconfirmation strategies when reading belief-inconsistent text. These results suggest that relevance task instructions are able to alter memory for belief-relevant information, but not the way in which readers process belief-consistent and belief-inconsistent information. In the context of the MD-TRACE model, such externally provided specific reading goals or task instructions influence particular reading situations because they affect the task model and the particular circumstances. Metacognitive strategy trainings, however, focus on readers' internal resources, targeting readers' knowledge and self-regulation of cognitions (Flavell, 1976). It includes knowledge of the cognitive processes that underlie mental activities such as thinking and comprehension and strategies of self-regulation and self-monitoring to control the direction, intensity and persistence of these cognitive processes.

In general, comprehending and integrating information from multiple texts necessitates the coordination of multiple processes, which is likely to be affected by metacognitive processes. In the context of belief-relevant multiple texts, enhancing readers' knowledge on how they comprehend the different texts with conflicting information and on what processes are involved during comprehension might increase their ability to monitor and regulate relevant cognitive processes in a functional way. This approach could foster readers' critical thinking about conflicting multiple texts (Ku & Ho, 2010; Maier & Richter, 2014). The importance of metacognitive knowledge for belief-relevant multiple-text comprehension is supported by McCrudden and Barnes (2016) who investigated the myside bias in evaluation of belief-consistent and belief-inconsistent arguments. In their study, secondary students differed in the strength of the myside bias, some showing more objectivity than others. Using qualitative data from interviews, the authors investigated the strategic processing of more-objective and less-objective participants. They found that more-objective participants used normative criteria for an in-depth scrutiny of belief-consistent and belief-inconsistent arguments, whereas less-objective participants scrutinized only belief-inconsistent arguments by using normative criteria but also irrelevant validation criteria. The authors concluded that less-objective and more-objective participants differ in their individual conceptual and procedural metacognition, which causes the myside bias. A similar effect occurred in the processing data from Bohn-Gettler and McCrudden (2018), showing a disadvantage in the strategic processing of belief-inconsistent text material for undergraduates. It is important to note that these studies only investigated belief effects during strategic processing. However, as noted earlier, beliefs also affect comprehension of belief-relevant multiple texts during passive and non-strategic processing stages (Maier et al., 2018). Metacognitive knowledge about such non-strategic belief effects on comprehension might therefore also be important.

To the best of our knowledge, only one study to date has investigated the effectiveness of a metacognitive strategy training on the comprehension of belief relevant multiple texts (Maier & Richter, 2014). In this study, undergraduate university students in the training condition received knowledge about 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. The first strategy was especially meant to enhance readers' awareness of belief effects during non-strategic processing, that is, awareness of how readers tend to favor belief-consistent information in text comprehension (Step 1 of the Two-Step Model of Validation, Richter & Maier, 2017). This strategy of the metacognitive training was expected to lead to an initiation of more strategic validation processes for belief-inconsistent texts and additionally should increase the persistence to continue this effortful processing until a good comprehension for belief-inconsistent information is reached. As such, becoming aware of the influence of prior beliefs on processing might lead to an adaption of the standard of coherence which readers might set differently for belief-consistent and belief-inconsistent texts. The additionally provided metacognitive strategies in the study from Maier and Richter (2014) were expected to assist readers to successfully and strategically process belief-relevant material. As such, they focused on relevant activities and monitoring processes (i.e., ongoing control of task-processing regarding one's cognitive goal) that readers should engage in during Step 2 of the Two-Step Model of Validation.

In addition to varying whether participants received the metacognitive training or no training, Maier and Richter (2014) also varied the motivational circumstances for two of three training conditions by providing external feedback in a prior task that was either positive or negative. One additional training group received no feedback. The results revealed a text-belief consistency effect for participants who had received no training, showing greater comprehension of a belief-consistent compared to a belief-inconsistent text. These participants spent equal amount of time reading the two text types. Participants in the training conditions, however, spent more time reading the belief-inconsistent text, but their comprehension outcome still differed depending on their motivation. A text-belief consistency effect was found for participants in the training groups that received either none or negative external feedback. Participants in the training group with positive feedback (i.e., beneficial motivational circumstances), however, comprehended the belief-inconsistent text to a similar extent as the belief-consistent text. This result suggests that more time devoted to reading belief-inconsistent material is not sufficient to successfully understand such text types. More crucial is how belief-inconsistent material is processed (see also McCrudden & Barnes, 2016).

1.4 Rationale and Overview

Previous research indicates that readers often have a one-sided mental representation of multiple texts in which belief-consistent information is represented to a greater extent (text-belief consistency effect, for an overview see Richter & Maier, 2017). Such a bias reduces readers' ability to fully understand controversially debated scientific issues, to critically weigh the evidence and to make informed judgements about the issue. Accordingly, assisting readers in the comprehension of multiple belief-consistent and belief-inconsistent texts is an important step in counteracting this bias. There is evidence that knowledge about metacognitive strategies on non-strategic and strategic processes involved in multiple text comprehension can reduce the text-belief consistency effect by increasing participants' comprehension of belief-inconsistent information (Maier & Richter, 2014). However, this effect only occurred when information about the metacognitive strategies was combined with beneficial motivational circumstances. Given that learning with multiple texts often involves self-regulated learning, constructing beneficial motivational circumstances externally is likely to be unreliable. Instead, a metacognitive training should be effective on its own. In two experiments we investigated whether training readers how to use metacognitive strategies focusing on belief-biasing validation processes during multiple text comprehension can reduce the text-belief consistency effect independently of motivational circumstances.

In the Maier and Richter study, the effectiveness of the metacognitive strategy knowledge intervention was not compared with the effectiveness of an alternative intervention. To address this research gap, we compared the effectiveness of a metacognitive strategy training for multiple text comprehension to the effectiveness of the PQ4R training (Thomas & Robinson, 1972) in the present studies. The PQ4R training is a well-situated and effective reading skill training that has been shown to increase elaborative processing (Thomas & Robinson, 1972) and has a long history of use (for an overview, see Slavin, 2014).

One explanation for the relevance of beneficial motivational circumstances in the study from Maier and Richter (2014) is that mere declarative prescriptions of metacognitive knowledge were used (e.g., the biasing influence of prior beliefs). In this respect, metacognitive knowledge comprises person-related knowledge (i.e., perceiving oneself as a thinker), task-related knowledge (i.e., recognizing the tasks' characteristics and demands), and knowledge about strategies and their effectiveness in coping with a particular task. Maier and Richter provided such information to undergraduate university students. In detail, knowledge about the person (e.g., the influence of validation and prior beliefs), about the task (e.g., information on the challenges readers encounter during the comprehension of belief-relevant multiple texts such as relating information across texts) and about activities of self-regulation and self-monitoring were provided. Metacognition, however, also includes regulation of cognitions (Flavell, 1976), which was not taught or practiced in the study from Maier and Richter. In the present experiments, we combined metacognitive knowledge with two elements that were viewed as relevant for the regulation of cognition during the comprehension of belief-relevant multiple text: observational learning and individual practice. In detail, the metacognitive training in our experiments provided participants with information on how to regulate their cognitions during the comprehension of belief-relevant multiple texts, that is, by allowing insight into self-regulation and self-monitoring of the cognitive processes in focus of the trainings. For this aim, learning from modeling examples (van Gog & Rummel, 2010) was integrated by watching role models who successfully or unsuccessfully applied the metacognitive strategies (or PQ4R steps in the control condition). This component of the training was further supported by either (a) practicing the use of the strategies/steps without external control or feedback (Experiment 1) and (b) noticing and explaining when someone uses the strategies/steps with elaborative feedback on the answers as well as testing factual knowledge about the strategies/steps (Experiment 2).

In both experiments, we expected the metacognitive strategy training to reduce the text-belief consistency effect compared to the PQ4R training. As such, the goal of the training was not to foster comprehension overall, but to reduce the gap between the comprehension of belief-consistent and belief-inconsistent texts by increasing the understanding for belief-inconsistent information. This is crucial because it should enable students to also understand valuable and reasonable counterarguments and to construct a comprehensive understanding of scientific issues.

2. Experiment 1

In Experiment 1, the metacognitive strategy training and the PQ4R training were provided in a group setting and consisted of a knowledge provision phase, an observational learning phase, and a practice phase. The metacognitive strategy training targeted five metacognitive strategies that were assumed to be important for the comprehension of multiple belief-relevant texts. In particular, we expected the five metacognitive strategies to reduce the comprehension disadvantage for belief-inconsistent texts. Three metacognitive strategies (becoming aware of the influence of prior beliefs, monitoring for intertextual relationships and inconsistencies, using prior knowledge for argument evaluation) have been shown to augment multiple text comprehension in an earlier experiment (Maier

& Richter, 2014). In addition, the metacognitive strategies a) monitoring for intratextual inconsistencies and b) memorizing facts were included in the training because independent research suggests that these strategies are required for successful multiple text comprehension (cf. Bråten et al., in press). The control condition was based on the PQ4R training (Thomas & Robinson, 1972). Participants in the control condition received knowledge about the six PQ4R steps (Preview, Questions, Read, Reflect, Recite, Review) and were trained on the application of these strategies in multiple text comprehension. In detail, the PQ4R training focuses on elaborative reading skills that are assumed to augment memorization and comprehension in six consecutive steps while reading a text (Thomas & Robinson, 1972). As such, the PQ4R training supports reading by providing learning strategies, but no metacognitive knowledge is provided, that is, no information on thinking about thinking. The main principle is that readers become actively engaged with the topic by (1) organizing the comprehension process (i.e., getting an overview of the structure and topic of the texts), (2) asking questions prior to reading the texts, (3) reading the texts with the questions in mind, (4) reflecting about the content of a particular paragraph and their understanding, (5) reciting the content of greater passages by paraphrasing it, and (6) reviewing the texts by comparing the answers to the questions about the text content.

We expected the metacognitive training but not the PQ4R training to prepare readers for the challenges of multiple text comprehension in the context of controversially discussed issues. The PQ4R training fosters the application of strategies that increase the receptive processing of information, such as the elaboration and integration of information, but not epistemic processing, which involves a critical evaluation of the epistemic status of information (Richter & Schmid, 2010; Münchow, Richter, & Schmid, in press). Therefore, the PQ4R training might help in detecting intertextual inconsistencies and inconsistencies between readers' prior beliefs and textual information but not in resolving these inconsistencies by increasing Step 2 processing. In other words, prior knowledge activation and elaboration, which are the focus of the PQ4R training, are not sufficient to assist readers in the successful comprehension of belief-consistent and belief-inconsistent controversial science texts. The strategies learned in the metacognitive training, however, are expected to support readers by especially augmenting their comprehension of belief-inconsistent information. Based on this assumption, we predicted a reduction of the text-belief consistency effect in the metacognitive strategic training but no such effect in the PQ4R training condition.

2.1 Method

2.1.1 Participants. Thirty-nine psychology students (29 women and 10 men) with an average age of 23.59 years (*SD* = 6.62) and an average semester of 2.77 (*SD* = 1.35) participated in the study. At the beginning of the experiment proper, participants' prior beliefs about the topic (i.e., risks and benefits of vaccinations) were assessed with two items (ratings on a scale from 0 = *do not agree* to 6 = *fully agree*). One item ("I believe that regular vaccinations are more beneficial than harmful") assessed participants agreement to the argumentative stance of the pro text that argued for the necessity and utility of vaccinations, and one item ("I believe that regular vaccinations are more harmful than beneficial") assessed participants' agreement to the contra argumentative text stance that vaccinations are unnecessary and harmful. A difference score (Item pro – Item con) was computed, with positive values indicating stronger agreement with the pro-vaccination text and negative values indicating stronger agreement with the contra-vaccination text. Overall, the mean difference score was positive (*M* = 2.77, *SD* = 1.90) indicating that participants more strongly agreed with the pro text. However, after scrutinizing the difference score for each participant, two participants were identified with a negative difference score, indicating a preference for the contra-text stance. Given the uneven cell sizes caused by the imbalance of belief-consistency, the data from participants that more strongly agreed with the contra argumentative text stance were not used in further analyses. In the remaining sample, 13.5% of participants (*n* = 5) had no clear preference (difference score of 0), 21.6% had low to medium belief strengths (*n* = 8, difference score of 1 to 2) and the remaining 64.9% (*n* = 24) had medium to strong beliefs (difference score of 3 to 5). In the following parts of the paper, the argumentative stance of the pro text is referred to as *belief-consistent position*, whereas the argumentative stance of the contra text is referred to as *belief-inconsistent position*.

The remaining sample consisted of 37 psychology undergraduate students (27 females, 10 males) with an average age of 24.14 years (*SD* = 5.49) and a mean study duration of 2.70 semesters (*SD* = 1.27). Participants received course credit for participating. From these 37 participants, 18 received the metacognitive training and 19 received the PQ4R training.

2.1.2 Trainings. Both training conditions followed the same structure and were provided in a group setting. The trainings started with a section on *knowledge provision* that used videos for observational learning and verbal explanations. In this section, participants first watched two short video scenes for each metacognitive strategy or for each PQ4R step. Participants were told that the videos showed other university students that were asked to express what they were thinking while reading two controversial texts on global warming. For each metacognitive strategy and for each PQ4R step, two role models were displayed—one role model playing a critical, sophisticated student who successfully uses the strategy or step, and one uncritical role model who is unable to successfully use the strategy or step. The effective or ineffective use of strategies and steps was developed by using preconstructed text modules that the role models spoke realistically during the recording (learning from modeling examples, van Gog & Rummel, 2010). In sum, ten video scenes (i.e., one critical and one uncritical role model scene for each of the five strategies) were viewed in the metacognitive strategy training condition and twelve video scenes (i.e., one critical and one uncritical role model scene for each of the six PQ4R steps) were viewed by participants in the PQ4R training condition. For example, the critical role model made the following comment for the metacognitive strategy *monitoring for intertextual relationships and inconsistencies*: “The texts are inconsistent with one another. The first text claimed that mankind caused global warming, whereas this text does state that natural circumstances caused global warming.” The uncritical role model for the same strategy stated, “The other text might have stated something about that, too.” Similarly, for the PQ4R step “questions,” one scene with a critical role model represented a valid and sound question that a reader might have for the texts (e.g., “What might be the cause for chances in global temperature?”), and one video scene with a different uncritical role model represented a reader who was not able to come up with a good question for the texts (e.g., “I will summarize this part. I cannot think of a question.”). After participants had watched the two videos for either a metacognitive strategy or PQ4R step, they received written and verbal information on each strategy or step, which explained its importance and how to use it in detail.

The next part of the training was *practice*. Participants were instructed to practice the use of the strategies or steps alone. For this aim, participants received an outline that summarized the strategies or steps and were instructed to work on two texts with opposing stances on the issue of man-made causes of global warming (pro text: 323 words; contra text: 328 words). Their task was to read the texts with the strategies or steps in mind and to find examples in the text when they should be applied. During this task, participants worked on their own and no explicit instruction was given on how to find or mark examples in the text. The goal of the practice phase was to provide participants with a better understanding of the metacognitive strategies or PQ4R steps and to increase their awareness of when and how to apply them when reading controversial science texts.

2.1.3 Text material. The experimental text material consisted of two texts discussing the risks and benefits of vaccinations and was based on text material used in a study by Maier and Richter (2014). One text argued that vaccinations are necessary and beneficial (pro-text stance) and the other text argued that vaccinations are unnecessary and harmful (contra-text stance). The texts were similar in all other respects such as length (average length: 899 words), writing style, structure, and readability (determined with the German adaption of the Flesch's Reading Ease Index, Amstad, 1978), as well as understandability, plausibility, interest, and number of arguments (for details, see Maier & Richter, 2014).

2.1.4 Comprehension measure. A verification task (adapted from Schmalhofer & Glavanov, 1986) was used to measure readers' situation model for each text. In the verification task, participants decided whether or not three different types of test items (paraphrases, inferences, distracters, eight items per item type and text) matched the situation described by the texts (examples of the test items are available in Table 1). For paraphrase items, an original sentence of the text was modified by changing the word order of the original sentence and replacing the key content word with synonymous expressions. Hence, paraphrase items contained information that was explicitly provided by the texts. In contrast, inference items contained information not presented in the texts but necessary for participants to build an adequate mental representation of the text content. Distracter items were only loosely associated with the text content by providing additional information about vaccination. The measure for each text situation model was based on a comprehension score that was corrected for response tendencies. For this aim, the proportions of the probit-transformed yes responses to distracter items (false alarms) were subtracted from the proportions of the probit-transformed yes responses to inference items (hits) (see Maier & Richter, 2014).

2.2 Procedure

The experiment proper consisted of a *training phase* (divided into observational learning, knowledge provision, and practice) and a *test phase*. At the beginning of the *training phase*, participants read the two texts on global warming to familiarize them with the topic, which was used in the video scenes. Next, half of the participants received the metacognitive strategy training and the other half the PQ4R training (for details on the trainings, see section 2.1.2). The training (including verbal explanations, watching the video scenes, and practicing) lasted 30-45 minutes. After the training and in the beginning of the *test phase*, participants' beliefs about vaccinations were assessed. Participants subsequently read the two texts on vaccination in a self-paced fashion, one paragraph at a time on a computer screen. A computer-based presentation of text material was used to align the experimental situation to the scenario of informal reading of web-based texts. Paragraph per paragraph presentation was used to avoid the need to scroll up and down to read the texts. The outlines that were given to participants during the training phase were available during reading and were removed after reading. After reading the two texts, participants worked on the verification task. The test items 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. Participants indicated for each test item whether or not the test item matched the situation that was described in the texts. Participants provided their judgments by pressing one of two response keys (marked green for yes and red for no). At the end of the experiment, participants were thanked and debriefed. The whole experimental session took about 1.5 hours.

2.3 Design

The experimental design was a 2 (*text-belief consistency*: consistent vs. inconsistent; varied within participants) x 2 (*type of training*: metacognitive training vs. PQ4R training, varied between participants) design. The text order (consistent-inconsistent vs. inconsistent-consistent) was counterbalanced between participants and included as a control factor.

2.4 Results and Discussion

The hypotheses pertaining to the effects of text-belief consistency and training were tested with an ANOVA for designs with between- and within-subjects factors. 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 .84 for detecting the focal interaction of text-belief consistency and training 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 2.

The ANOVA revealed an interaction of text-belief consistency and training, $F(1, 33) = 5.30$, $p < .05$, $\eta_p = .14$ (Figure 1). As expected, participants that had received the PQ4R training had a better situation model for the belief-consistent text ($M = 1.81$, $SE_M = 0.16$) compared to the belief-inconsistent text ($M = 1.43$, $SE_M = 0.15$), $F(1, 33) = 4.94$, $p < .05$, $\eta_p = .13$. In contrast and also as expected, for participants in the metacognitive strategy training condition, no difference was found in situation model strength between the belief-consistent ($M = 1.14$, $SE_M = 0.17$) and the belief-inconsistent texts ($M = 1.32$, $SE_M = 0.15$), $F(1, 33) = 1.10$, $p = .30$, *n.s*. Moreover, we also found a main effect of the training, $F(1, 33) = 4.26$, $p < .05$, $\eta_p = .11$. Participants' situation model was overall better for participants who had received the PQ4R training ($M = 1.62$, $SE_M = 0.13$) compared to participants who had received the metacognitive strategy training ($M = 1.23$, $SE_M = 0.14$). No other effects of the independent variables were significant.

The results of Experiment 1 were in line with the predictions. The strategies learned in the metacognitive training condition helped readers to reduce the text-belief consistency effect in multiple text comprehension. The PQ4R training, however, was not sufficient to reduce the text-belief consistency effect in comprehension. Instead, participants in this condition had a stronger situation model of the belief-consistent text. These results support the assumption that a training of metacognitive knowledge and practice on regulation can reduce the text-belief consistency effect without additional beneficial motivational circumstances. However, the balanced situation models for belief-consistent and belief-inconsistent texts in the metacognitive strategy condition came at the cost of slightly reduced situation models for the two texts, especially for the belief-consistent text. A similar pattern of results had been found in earlier research, for example, in a study investigating multiple text comprehension with an alternating presentation of belief-consistent and belief-inconsistent texts (Maier & Richter, 2013; Maier et al., 2018). This pattern of results across three studies suggests that readers might need more support to be able to fully comprehend both text types.

3. Experiment 2

The metacognitive strategy intervention used in Experiment 1 reduced the text-belief consistency effect in readers' understanding of a scientific controversy. However, the reduction of the text-belief consistency effect came at the cost of a reduced situation model for the belief-consistent text. This pattern of effects might have resulted from the rather minimalistic training used in Experiment 1, which created a condition of uncertainty and high mental effort needed to use the metacognitive strategies. Put differently, the pattern of results suggests that the intervention used in Experiment 1 might not have been comprehensive enough to induce strong comprehension of both text types. In Experiment 1, the training consisted of observational learning, knowledge provision, and a short practice trial, all of which are core elements of a successful training. In the observational phase, participants watched role models (un-)successfully applying the strategies or steps during reading. However, the video scenes were limited to prototype responses consisting of a few sentences when, for example, paying attention to prior beliefs. In addition, participants in the practice phase of Experiment 1 worked independently in finding examples of how to apply the trained content to multiple texts. Participants' responses during this phase were not coded or analyzed. Hence, we cannot be sure whether participants were actually successful in this practice trail and how they indicated or justified the application of the metacognitive strategies or the PQ4R steps in the control training. From our point of view, these shortcomings of Experiment 1 could have contributed to participants having misinterpretated the first metacognitive strategy, that is, becoming aware of the influence of prior beliefs. Participants might have had the impression that this strategy required them to comprehend the belief-inconsistent text better than the belief-consistent text. This interpretation is consistent with the findings from Bohn-Gettler and McCrudden (2018) showing that the task instruction that readers are following during reading influences their memory for texts. Hence, participants in the metacognitive strategy condition of Experiment 1 might have placed the belief-consistent text aside, focusing mainly on the belief-inconsistent text. Given that a relevance instruction does not necessarily change the way readers process the information (Bohn-Gettler & McCrudden, 2018), this focus might have balanced out the understanding for both texts but at the cost of a lower overall understanding.

Experiment 1 also suffered from additional limitations. First, the steps of the PQ4R training were presented in separate video scenes in the observational learning block to provide participants in the two training conditions with a similar number of video clips. However, given that the PQ4R steps are interdependent, a more reasonable approach would be to provide participants in the current experiment with all the steps together.

Second, to keep the length of both trainings similar (see method section for further details) and within reasonable time limits, we limited the metacognitive strategies in Experiment 2 to the three that have already been shown to be relevant for multiple text comprehension (Maier & Richter, 2014). These strategies were (1) becoming aware of the influence of prior beliefs, (2) monitoring for intertextual inconsistencies, and (3) using prior knowledge for argument evaluation. Whereas strategy 1 was expected to increase readers' awareness for the unconscious preference of belief-consistent information in text comprehension (Step 1 of the Two-Step Model of Validation, Richter & Maier, 2017), the two additional strategies provided information on processes considered relevant for Step 2 of the Two-Step Model of Validation, that is, paying attention to conflicting information and scrutinizing information based on prior knowledge (not beliefs). In addition, returning to the three metacognitive strategies investigated in earlier research allowed us to spend more time on training each of these strategies.

Third, the use of the strategies or steps was practiced independently by participants in Experiment 1 without any external control or feedback. In Experiment 2, we standardized the practice phase (for details see section 3.1.2). In addition, we decided to add elaborative feedback as an element to both training conditions because feedback has reliably been shown to be an effective mean to enhance learning from text (see, e.g., the meta-analysis by Swart et al., 2019; Hattie & Timperley, 2007). We were expecting that elaborative feedback should increase participants' understanding of the training content and how they should use the metacognitive strategies or PQ4R steps successfully during reading.

Fourth, Experiment 1 provided the training within a group session and with verbal explanations given by a trained expert on the topic. In Experiment 2, the whole experiment was computer-based. This allowed us to provide automated feedback to participants at various points during the training. Furthermore, it gave participants the possibility to work at their own pace and to reread the theoretical explanations of the training content if they encountered comprehension difficulties. In addition, the generated data from the training (e.g., training duration or the amount of correct responses in the knowledge application block, see section 3.1.2) allowed us to investigate whether or not participants understood the content of the trainings correctly and were able to identify how to use the strategies or steps when reading correctly.

Fifth, only one scientific topic was used in Experiment 1 for the experimental text material. We used two scientific controversies as text material in the second experiment to obtain information about the generalizability of the empirical findings and the strategies conveyed by the trainings.

Sixth, the results of Experiment 1 suggest that comprehending multiple texts while applying unfamiliar metacognitive strategies to the reading context is a challenging task for readers. The extent to which readers engage in such demanding cognitive tasks may systematically influence their comprehension processes and outcomes. Hence, successfully comprehending multiple texts and also successfully profiting from the training might depend on reader characteristics. Earlier research has provided evidence that individual differences with regard to epistemological beliefs (Mason & Boscolo,2004), belief strength (Kardash & Scholes, 1996; McCrudden & Barnes, 2016), prior knowledge, and level of education (Wiley, 2005) can influence the comprehension of belief-relevant multiple text (for a detailed discussion see Richter & Maier, 2017). In Experiment 2, we investigated weather readers' disposition to feel joy when thinking about effortful cognitive problems (i.e., their need for cognition; Cacioppo & Petty, 1982) influences the comprehension of belief-relevant multiple texts and also the potential beneficial effects of the trainings. Need for cognition has been shown to be strongly associated with better reading comprehension of both narrative and expository texts (Dai & Wang, 2007). Integrating conflicting argumentative stances into one coherent mental representation of the controversy can be viewed as a very complex task that might also depend on readers' need for cognition (Richter, 2011). In line with this assumption, Kardash and Scholes (1996) found that undergraduate students with higher need for cognition were more likely to draw differentiated conclusions that reflected the two-sidedness of the topic after reading controversial (scientific) information (i.e., "Does HIV cause AIDS?"). The authors suggested that readers with higher need for cognition have a tendency to more deeply scrutinize (conflicting) information. Given that the present study aimed at investigating the effectiveness of a metacognitive training, regardless of external motivational influences, operationalizing need for cognition as intrinsic (cognitive) motivation is reasonable. Moreover, investigating the interplay of readers' cognitive motivation and a metacognitive training may extend the understanding of how interventions and individual characteristics interact during the comprehension of belief-relevant multiple texts.

Similar to Experiment 1, we predicted that participants in the metacognitive strategic training should comprehend belief-consistent and belief-inconsistent texts equally well, whereas participants in the PQ4R training condition should show a text-belief consistency effect in comprehension, that is, a better situation model for a belief-consistent compared to a belief-inconsistent text (Hypothesis 1). Moreover, need for cognition was expected to reduce the text-belief consistency effect (Hypothesis 2a) and participants with a higher need for cognition were expected to especially benefit from the metacognitive strategy training (Hypothesis 2b). In particular, applying unfamiliar metacognitive processes to a complex reading scenario requires readers to think about and regulate their cognitions. Using the learning strategies proposed in the PQ4R steps might be less challenging and thus might depend less on readers' need for cognition as intrinsic (cognitive) motivation.

3.1 Method

3.1.1 Participants. Fifty-three psychology students (39 women and 14 men) with an average age of 22.0 years (*SD* = 5.70) participated in the training study. All participants were rewarded with course credit.

Similar to Experiment 1, we first investigated participants' prior beliefs about the two text' topics (i.e., training of intelligence and vaccinations). In Experiment 2, ten items each were used to assess participants' prior beliefs on the two scientific issues of intelligence and vaccinations (ratings on a scale from 0 = *do not agree* to 6 = *fully agree*). Participants responded to five statements that assessed their agreement with the view that vaccinations are more beneficial than harmful (e.g., "From my point of view, vaccinations are the most important and most effective mean against infectious diseases", Cronbach's $\alpha = .81$) and five statements that assessed their agreement with the argumentative stance that vaccinations are more harmful than beneficial (e.g., "I think that vaccinations are more harmful than beneficial", Cronbach's $\alpha =.91$). Similarly, five statements assessed participants' beliefs that intelligence can be trained (e.g., "I think that cognitive training can be used to booster intelligence", Cronbach's $\alpha =.86$) and another five assessed participants' beliefs that intelligence cannot be training and is rather innate (e.g., "I think that intelligence is innate", Cronbach's $\alpha =.60$). For both topics, mean agreement to the four belief scales were estimated.

Participants overall agreed more with the pro argumentative stance in the two scientific issues. In detail, participants agreed more with the belief that vaccinations are beneficial (*M* = 3.44, *SD* = 1.02) and less with the belief that vaccinations are harmful (*M* = 2.07, *SD* = 1.04), $t(52) = 5.22$, $p < .05$, and more strongly endorsed the view that intelligence can be trained (*M* = 4.43, *SD* = 0.98) and agreed less with the proposition that intelligence cannot be trained (*M* = 3.25, *SD* = 0.73), $t(52) = 6.10$, $p < .05$. Nevertheless, we also identified three participants for each topic condition that agreed more with the contra argumentative stance in the two issues. Following the procedure from Experiment 1, data from these participants were omitted from further analyses. In the remaining sample, 4.3% of participants (*n* = 2) had no clear preference (difference score of 0), 61% had low to medium belief strengths (*n* = 28, 0 < difference score ≤ 2) and 35% (*n* = 16) had medium to strong beliefs (difference score > 2). The argumentative stance of the pro texts is referred to as *belief-consistent position*, and the argumentative stance of the contra texts is referred to as *belief-inconsistent* position.

The remaining sample consisted of 46 first semester psychology students (35 females, 11 males) with an average age of 21.20 years (*SD* = 4.34). [1] Participants received course credit for their participation in the study. From these 46 participants, 25 participants were in the metacognitive training group and 21 in the PQ4R training group.[2]

3.1.2 Trainings. The two training interventions (metacognitive vs. PQ4R) were implemented as computer-based trainings with Inquisit (Version 4.0.8, 2015). Participants worked independently and at their own pace. Both trainings followed a similar structure that consisted of 1) a theoretical

[1] Data of one additional participant was excluded from the analyses as this student was studying in his 10th semester, whereas all other participants were in the first semester.

[2] The difference between the number of participants in the training groups is due to excluding participants that more strongly agreed with the contra argumentative text stance.

explanation, 2) observational learning, 3) knowledge application with feedback, and 4) knowledge check (see Figure 2).

In the *theoretical explanation block,* participants received either detailed written explanations for the metacognitive strategies (for the exact wording translated into English see Table 3) or the PQ4R steps. Participants in the metacognitive strategy training received a theoretical explanation for one strategy at a time, whereas explanations for all PQ4R steps (on separate pages) were given at this point in the PQ4R training condition. Explaining all six steps at one time made sense because of the interdependence among the steps.

In the *observational learning block,* participants watched videos that displayed a reader who was either ineffectively or effectively using the metacognitive strategy or PQ4R steps by thinking aloud while reading two controversial texts on cancer check-ups. In the metacognitive strategy training condition, participants received a positive (i.e., a critical role model effectively applying a metacognitive strategy) and a negative (i.e., an uncritical role model ineffectively applying a metacognitive strategy) example video for each strategy. The order of positive and negative videos was varied between the metacognitive strategies. Six videos were watched in total in the metacognitive strategy intervention, whereas only two video examples were watched in total in the PQ4R training because of the interdependence among the steps. In detail, participants in the PQ4R training condition first watched a video example of a reader who only applied two PQ4R steps effectively and four steps ineffectively (negative example) and afterwards watched a positive video example in which a role model successfully applied all PQ4R steps during multiple text comprehension.

In the *knowledge application with feedback block*, participants in both training conditions received elaborative feedback. During this phase of the trainings, participants received feedback twice. First, participants answered single-choice tasks with four-answer options (one correct statement, two incorrect statements and “do not know” option) on whether the role model in the video from the observational learning block successfully applied or failed to apply the respective metacognitive strategy (e.g., for the first metacognitive strategy: “Did the student in the video become aware of his or her prior beliefs and their impact while reading controversial texts?’) or all PQ4R steps (e.g., “Did the student in the video take all of the PQ4R steps into account while reading controversial texts?”). Participants were additionally asked to justify their answer in one or two sentences. After giving their responses, participants were provided with feedback about the accuracy of their response in the single-choice question and also with the correct response. This knowledge of correct response (i.e., student model did/did not use the metacognitive strategy/PQ4R steps in the video) was then further accompanied by an explanation when and how the strategy/step was used/not used by the role models by providing characteristic scenes from the video with a written description explaining how the role model was successful/unsuccessful in using the strategy/steps. In the metacognitive training condition, two characteristic scenes from the video of the observational learning block were presented for each strategy that were designed to clarify the ineffective or effective use of the previously explained strategy. In the PQ4R training condition, participants watched six characteristic scenes from the video of the observational learning block that were designed to clarify the manner in which the person in the video dealt with each PQ4R step.

Finally, in the *knowledge check block*, participants’ knowledge of the metacognitive strategies or PQ4R steps was assessed using two single-choice tasks with four-answer options for each strategy or step (one correct statement, two incorrect statements and “do not know” option). In the metacognitive strategy training, participants answered six knowledge check questions (two for each of the three metacognitive strategies). In the PQ4R condition, participants answered twelve knowledge check questions (two for each of the six PQ4R steps). Participants needed to mark the statement that applied to each strategy or step in the first question. In the second question, participants needed to mark the statement that does not apply to the strategy or step. Participants who failed to answer the questions correctly the first time had the opportunity to reread the description of the strategies/steps. If such participants failed to answer the question correctly the second time, they were provided with knowledge of the correct response. Participants who answered the factual single choice questions correctly the first or second time, received positive feedback (i.e., “great, your answer was correct”). Feedback was based on an automatic coding of participants’ responses in the single choice questions, was delivered by the computer, and given individually. After answering all questions, the training was completed.

3.1.3 Text material. The experimental text material consisted of the two texts that were used in Experiment 1 (usefulness of vaccinations) and two texts that addressed the possibility to train one’s intelligence. For the new scientific topic, the first text argued for the possibility of training one’s intelligence (pro-text stance), whereas the second text argued that intelligence is hereditary and cannot be trained (contra-text stance). The texts for the topic intelligence were similar in all other respects such as length, writing style, structure, and readability (determined with the German adaption of the Flesch’s Reading Ease Index, Amstad, 1978), as well as understandability, plausibility, interest, and number of arguments (Table 4). Paired-samples *t* tests (Holm-Bonferroni correction for multiple tests, Holm, 1979) revealed no significant differences between the texts in any of the text characteristics in a pilot test with an independent sample of university students.

3.1.4 Comprehension measure. As in Experiment 1, a verification task (adapted from Schmalhofer & Glavanov, 1986) was used to measure readers’ comprehension of each text on the level of the situation model (Table 1).

3.1.5 Need for cognition. Need for cognition was assessed with the German version of the need for cognition scale (Bless, Wänke, Bohner, Fellhauer, & Schwarz, 1991) that was originally published by Cacioppo and Petty (1982). Participants rated 16 items on the degree to which they enjoy cognitively demanding tasks and problems (response categories ranged from 1 = *strongly disagree* to 7 = *strongly agree*, Cronbach's α = .90).

3.2 Procedure

Approximately two weeks before to the main experiment, participants' prior beliefs and their need for cognition were assessed online. In the main experiment, participants were randomly assigned to the two training conditions (metacognitive strategies vs. PQ4R). The main experiment was fully presented on computer. Before starting with the training, participants read the two texts about the risks and benefits of cancer check-ups that were used as text material in the training. The procedure of the computerized trainings was strictly parallelized (see the section 3.1.2). After the training, participants read one belief-consistent and one belief-inconsistent texts on one of the two scientific issues (training of intelligence vs. vaccinations). Similar to Experiment 1, texts were presented on computer screen paragraph by paragraph. Participants were able to freely navigate between paragraphs during reading. After reading, participants worked on the verification task. After completing all tasks, participants were thanked and debriefed. The main experiment was 120 min long, with the trainings lasting approximately 60 min (including the videos).

3.3 Design

The design was a 2 (*text-belief consistency*: belief-consistent vs. belief-inconsistent, varied within subjects) x 2 (*type of training*: metacognitive vs. PQ4R, varied between subjects) design. Need for cognition was included as a covariate. Text order (consistent-inconsistent vs. inconsistent-consistent) and text topic (training of intelligence vs. vaccinations) were counterbalanced between participants to control for ordering and topic effects.

3.4 Results and Discussion

3.4.1 Descriptive Data for the Training.

On average, participants answered 80% of the knowledge check questions correctly (M = .79, SD = .15). The range of wrong answers was 0 to 4 in the PQ4R condition and 0 to 3 in the metacognitive strategy condition. We also investigated the extent that participants reread the explanations of the strategies/steps. This option was only used by four participants in the metacognitive strategy training condition and by nine participants in the PQ4R condition.

3.4.2 Effects of Beliefs and Trainings on Comprehension. Similar to Experiment 1, the hypotheses pertaining to effects of text-belief consistency and training were tested with an ANCOVA for designs with between- and within-subjects factors including reading order and text topic as control factors and need for cognition (*z*-standardized) as covariate. All hypothesis tests were based on Type I error probability of .05. The power (1-β) for detecting the focal interaction of text-belief consistency and training condition was .91 (computed with G*Power 3; Faul et al., 2007). Descriptive statistics and intercorrelations of all variables are provided in Table 5.

We found a main effect of text-belief consistency in the verification task indicating that, overall, participants situation model for the belief-consistent text (M = 2.20, SE = 0.12) was stronger compared to participants situation model for the belief-inconsistent texts (M = 1.87, SE = 0.11), $F(1, 37) = 7.7$, $p < .05$, $\eta_p^2 = .17$. However, this main effect was qualified by an interaction of text-belief consistency and training, $F(1, 37) = 4.0$, $p = .05$, $\eta_p^2 = .10$ (Figure 3). Participants receiving the PQ4R training had a stronger situation model for the belief-consistent text (M = 2.28, SE = 0.17) compared to the belief-inconsistent text (M = 1.72, SE = 0.16), $F(1, 37) = 10.6$, $p < .05$, $\eta_p^2 = .22$. In contrast, no difference was found in the situation models for the belief-consistent (M = 2.12, SE = 0.16) and the belief-inconsistent text (M = 2.03, SE = 0.14) in the metacognitive training condition, $F < 1.0$, $p = .54$, *n.s.*. This pattern of results fully supported Hypothesis 1.

We also found a main effect of need for cognition, $F(1, 37) = 9.1$, $p < .05$, $\eta_p^2 = .20$. Need for cognition and participants' performance in the verification task were positively correlated ($r = .48$, $p <$

.05), indicating that participants with higher values of need for cognition performed better in the verification task. This result is in line with Hypothesis 2a and earlier research showing that readers who like to take cognitive challenges outperform readers that do not enjoy cognitive thinking. We found no significant differential effect of the two trainings for participants with higher or lower need for cognition, indicated by the nonsignificant interaction of need for cognition and training, $F(1, 37) = 2.9$, $p = .10$.

4. General Discussion

In two experiments, we examined the effectiveness of a metacognitive strategy training on its ability to reduce the text-belief consistency effect in multiple text comprehension. Results from both experiments revealed that participants trained with the metacognitive strategy training comprehended belief-consistent and belief-inconsistent texts equally well, whereas participants trained with the control training based on the PQ4R steps showed a text-belief consistency effect in comprehension.

4.1. Theoretical Implications

Four main conclusions can be drawn from the results. First, the observed text-belief consistency effect for the PQ4R training condition is a replication of earlier findings (e.g., Abendroth & Richter, 2020a; Maier & Richter, 2013, 2014), lending further support to the assumption that readers' prior beliefs bear a strong impact on the comprehension of unfamiliar but belief-relevant scientific debates. These findings are of great practical significance because readers reading information online are regularly confronted with unfamiliar and complex scientific debates for which they lack the ability to fully understand and evaluate the presented context (e.g., Keil, 2010). Nevertheless, such information is often presented in articles on reputable and highly accessed online newspapers in a superficial way, which results in perpetuating readers' unawareness of their inability to fully make sense of the scientific discourse (Scharrer, Stadtler, & Bromme, 2014). In such circumstances, readers seem to rely on their prior beliefs as epistemic background to choose information for deeper processing (Richter & Maier, 2017; see also Maier, et al., 2018). Hence, finding mechanisms to reduce the text-belief consistency effect in multiple text comprehension is not an end in itself but an important step in increasing readers' awareness about the unsure nature of scientific debates and their ability to gain a full understanding of scientific disputes that are relevant for individual decisions.

Second, the results from the PQ4R training condition indicate that trainings focusing on receptive elaborative strategies are not sufficient when readers are trying to understand and evaluate complex scientific but belief-relevant multiple texts. In many reading scenarios, fostering text-based elaborations appears to be highly effective because the construction of a strong situation model requires the integration of textual information and readers' prior world knowledge (Kintsch, 1988). However, this approach is not sufficient for the comprehension of belief-relevant controversial texts. It might be expected that by activating prior knowledge, the detection of inconsistencies with prior knowledge becomes easier. Nevertheless, activating prior knowledge still does not provide an advantage for resolving such inconsistencies, which we consider a prerequisite for strategic validation processes. The PQ4R training might also have extended the passive and non-strategic influence of prior beliefs given that the formulated questions to the text as a core element of PQ4R might even be biased towards readers' beliefs (e.g. "How damaging are vaccinations?" vs. "What are the risks and benefits of vaccinations?"). In sum, an effective training for multiple text comprehension needs to activate all three cognitive processes involved in situation model construction, that is, the activation, integration, and validation of information (Richter, Münchow, & Abendroth, 2020).

Third, the effectiveness of the metacognitive strategy training compared to the PQ4R training in Experiment 2 clearly shows the importance of increasing readers' awareness of the belief-biasing validation processes that influence (multiple text) comprehension and the importance of training readers how to overcome such influences by strategic validation processes. This interpretation is consistent with the predictions of the Two-Step Model of Validation (Richter & Maier, 2017). In this respect, readers need to become aware that their prior beliefs operate as epistemic gatekeepers that bias the situation model construction towards belief-consistent information. This mechanism was the focus of the first strategy used in the metacognitive strategy training, which was expected to increase readers' self-awareness as a prerequisite to engage in self-regulation processes while learning (Baker & Brown, 1984). If readers understand that they often monitor the plausibility of new information by simply evaluating its consistency with prior beliefs, they can overcome this influence. The additional metacognitive strategies provided readers with further information of how to overcome the possible challenges of multiple text comprehension. In essence, these strategies provided advice on how to use prior knowledge—in contrast to prior beliefs—as a reliable source for evaluating the plausibility of new information. In terms of the MD-TRACE model (Rouet & Britt, 2011) and the RESOLV model (Rouet et al., 2017), the metacognitive strategies might have altered readers' initial task model. Readers' initial (explicit or implicit) understanding of the reading goal might have been to read for support of one's own beliefs. In the course of the metacognitive training, this default reading goal might have changed to the goal to read and understand multiple perspectives regardless of prior beliefs. Hence, all three of the metacognitive strategies could be viewed from the MD-TRACE model and the RESOLV model as strategies attached and used for specific goals. Note that in Experiment 1, the reduction of the text-belief consistency effect came at the cost of a reduced understanding of the belief-consistent text. In essence, we posit that participants had misinterpreted how to use the metacognitive strategies because no feedback was given on the practice phase in this study (for a full discussion, see section 2.4). In Experiment 2, however, the metacognitive training was effective in such a way that participants in this training condition were able to comprehend both text types equally well and to a similar extent as the belief-consistent text in the PQ4R training condition.

Fourth, the results provide evidence that the extended and computer-based metacognitive training of Experiment 2 consisting of knowledge provision, observational learning, external feedback, and individual practice is sufficient to overcome the text-belief consistency effect in comprehension on its own. This is an important extension of earlier research that showed only a beneficial effect of metacognitive knowledge when combined with performance feedback as a means to increase readers' motivation (Maier & Richter, 2014). The content of the metacognitive strategies tested in the present study was similar to the content provided by Maier & Richter (2014). However, the additional training elements used especially in Experiment 2 seem to be promising means for assisting readers in the comprehension of belief-relevant multiple texts. We postulate that observational learning and feedback on factual knowledge about the trained content and feedback on recognizing when a metacognitive strategy is successfully used during reading is crucial for the regulation of text processing. Based on the level of correct responses and the amount of rereading in the knowledge check block in Experiment 2, we conclude that overall participants understood the content of the trainings quite well. We did not, however, investigate individual paths within the trainings because our focus was on the comparison of the effectiveness of the two types of trainings, not on analyzing individual learning processes in the trainings. Nevertheless, investigating individual differences in broader and more diverse samples would be fruitful. For example, working memory capacity can be assessed to compare students who are challenged more by the training (i.e., have a higher amount of wrong answers and more re-readings of the explanations) and students who are challenged less by the training. Arguably, the cognitive demands of some students' multiple text comprehension could be already close to or even exceeding working memory capacity, which makes it less likely that they can profit from a demanding training during the first run. However, the training might be efficient for these students after more training sessions have been completed. In a longitudinal study, this could be investigated by aligning research data of the learning process of two participant groups that differ in individual working memory capacity to research data on the comprehension scores.

4.2. Limitations of the Present Research and Avenues for Future Research

We tested the combined effectiveness of the three metacognitive strategies because of their crucial interplay. In terms of the Two-Step Model of Validation (Richter & Maier, 2017) and the notion of standard of coherence (van den Broek, Beker, & Oudega, 2015), prior beliefs might lead to fairly low standards of coherence for belief-inconsistent texts during non-strategic validation (i.e., during regular comprehension, readers are not bothered by not understanding belief-inconsistent information or arguments). Hence, without conscious control, readers might set a different standard of coherence for belief-consistent information and belief-inconsistent information. From our point of view, the metacognitive training fundamentally enhances readers' awareness of the influence that prior beliefs can have on their comprehension of belief-relevant texts. Strategy 1 especially targeted Step 1 of the Two-Step Model of Validation and the automatic influence of prior beliefs on comprehension (i.e., non-strategic validation as a routine part of comprehension, Richter & Maier, 2017). Without making readers aware of how their beliefs influence comprehension, additional metacognitive strategies of how to handle belief-relevant multiple texts cannot be effective. Conversely, only making readers aware of the biasing influence of their prior beliefs is also insufficient. The additional trained metacognitive strategies were necessary to augment elaboration of the presented information from multiple texts, especially for the belief-inconsistent information. In part, the metacognitive training might also have led to higher goals or standards of coherence set for comprehension, that is, a reformulation of the task model in terms of the MD-TRACE Model (Rouet & Britt, 2011). As such, the use of relevant activities and monitoring processes (i.e., ongoing control of task processing regarding one's cognitive goal) that readers should engage in during Step 2 of the Two-Step Model of Validation might have been enhanced.

In the present study, the effectiveness of the metacognitive strategy training was investigated using university students. For the metacognitive strategy intervention to be effective, basic reading skills and self-regulation abilities to handle multiple texts are required, which limits the study population to older students. Further research should clarify if the metacognitive strategy training is also effective for other study populations such as adolescents in upper high school. Although adolescents possess necessary basic reading skills, they often lack relevant background knowledge for complex scientific controversies and also relevant metacognitive knowledge of how to deal with uncertainty and partial or even conflicting information. Consequently, adolescents have more difficulty comprehending conflicting science-related texts, and their prior beliefs have a strong influence on the comprehension of belief-relevant multiple texts (Abendroth & Richter, 2020a). We believe that investigating the effectiveness of the metacognitive strategy intervention for adolescents is an important direction of future research on this topic, but probably needs more training and practice sessions. The results of Experiment 1 suggest that a more basic training, which consists of the same elements such as observational learning, knowledge provision, and a short practice trial, but is reduced in size seems not comprehensive enough to induce strong comprehension of both texts. Rather, the successful application of the metacognitive strategies depends on an in-depth training with automated feedback that allows readers to seize and if necessary, correct one's understanding of the use of the metacognitive strategies. It seems likely that these training elements are even more important for adolescent readers. Research on this issue might address the effective use of the metacognitive strategy intervention embedded for example within the school curriculum and within classroom settings. Using more practice sessions on more controversial science-related topics that are part of school curriculum with feedback on comprehension performance and continuous follow-up tests within a school year might be necessary for adolescents to fully understand how to successfully use the metacognitive strategies. Similar research on university students over the course of several training sessions as part of the study program, might also allow for insights into the conditions that foster readers' spontaneous use of the strategies outside the training sessions.

In both experiments, text materials were presented on screen paragraph per paragraph. This presentation mode was used because training and reading situation were supposed to align to the reading scenario of web-based informal learning. The results suggest that the metacognitive strategy training was effective for the comprehension of belief-relevant multiple texts that are presented on screen. However, as reading texts on paper is suggested to be associated with an enhanced awareness of one's performance (Clinton, 2019), we consider it likely that the metacognitive strategy training is also effective when belief-relevant multiple texts would be presented on paper. That said, it is also possible that presenting the texts paragraph per paragraph on screen might have induced a more careful processing of the text material that might not occur during informal reading. However, the text-belief consistency effect found for the PQ4R condition suggest that the paragraph-per paragraph presentation alone was at least not sufficient to remediate the text-belief consistency effect. Additional research should clarify the effectiveness of the metacognitive strategy training on the comprehension of paper-presented multiple texts as well as for a computerized presentation of multiple texts that presents all textual information on one screen.

Experiment 1 and Experiment 2 did not use balanced samples in which participants preferring either of the two argumentative position on the two issues were equally represented. In choosing scientific topics for the experimental material, we were focusing on a strong manipulation of text-belief consistency taking into account that only a between-text variation of text-belief consistency seemed appropriate to reach this goal. Moreover, the effect of primary theoretical interest was the interaction effect of text-belief consistency with the two different trainings (metacognitive vs. PQ4R) with the latter factor varied experimentally as a between-participant factor. In addition, in the course of the material construction, a great amount of effort was taken to ensure that the text types were as similar as possible and showed no significant differences in objective characteristics for example with regard to understandability or plausibility. In future studies, it would be worthy to investigate the effectiveness of the metacognitive training on scientific topics for which balanced samples are possible, even if this might result in lower belief strengths.

One factor crucial for (multiple) text comprehension is prior topic knowledge (Braasch & Bråten, 2017), which is viewed as pre-requisite for elaborative processing to resolve inconsistencies between multiple texts (Richter & Maier, 2017). In the two training experiments presented here, prior knowledge had been assessed using multiple choice tests. However, based on rather low reliabilities (Cronbach's alpha) of these measures in Experiment 1 (Cronbach's alpha = .59, 12 Items) and Experiment 2 (vaccinations: Cronbach's alpha = .60, 7 Items, intelligence: Cronbach's alpha = .57, 12 Items), we decided not to include this measure in the data analyses. Previous research has revealed inconsistent findings on the influence of prior knowledge on the comprehension of belief-relevant multiple texts (e.g., Abendroth & Richter, 2020a; Wiley, 2005). In a study by Wiley (2005), in which the level of prior topic knowledge was varied using specific groups of university students, a text-belief consistency effect in argument recall was found for low-knowledge readers, but not for high-knowledge readers. Using prior knowledge as continuous covariate, Abendroth and Richter (2020a) found the text-belief consistency effect for adolescents with higher prior topic knowledge. It is important to note, however, that the overall level of prior knowledge in the study from Abendroth and Richter was quite low. In a recent study (Abendroth & Richter, 2020b), the effects of prior knowledge and prior beliefs were further disentangled. In this study, prior beliefs were experimentally varied using videos, which provided identical factual information but took opposing argumentative claims on a scientific issue. Results revealed an influence of such experimentally induced prior beliefs on the comprehension of multiple texts about an unfamiliar scientific topic. These results suggest that in line with assumptions made in the Two-Step Model of Validation, prior knowledge is not a prerequisite for text-belief consistency effects to occur. In addition, the metacognitive strategy intervention was directed at reducing belief effects and did not provide any content knowledge related to the texts. Nevertheless, it was effective reducing the text-belief consistency effect, whereas the PQ4R training focusing on the activation of prior knowledge was not. Nonetheless, additional research should be conducted to further investigate the relationship of prior knowledge and prior beliefs.

Understanding multiple texts often requires readers to build an intertext model and an integrated mental model. In the present study, we investigated the effectiveness of different trainings on the text-belief consistency effect, that is, on the comprehension outcome for belief-consistent and belief-inconsistent texts. From our point of view the comprehension of individual texts are one major determinant of the integrated mental model. In addition, the text-belief consistency effect is a well-established finding (Richter & Maier, 2017), allowing researchers to test hypotheses about training effectiveness. Nevertheless, our training studies have not investigated whether readers compared and contrasted the conflicting viewpoints from the multiple texts. Investigating the effects of the training on the integration of information across texts is also important and intertextual verification tasks or argument tasks requiring the integration of information seem suitable to investigate the integrated mental model. Given that one of the strategies trains readers on monitoring for intertextual inconsistencies and relationships, we assume that the metacognitive strategy training also enhances readers' ability to create a more integrated representation of the texts. This assumption, however, needs to be empirically tested. Moreover, we investigated no processing data in the present study. We recommend to align the present research findings with data from online processing such as eye-tracking or reading times in further research.

In the knowledge application with feedback block of Experiment 2, several components such as corrective feedback and also knowledge of correct response were included to support participants' understanding of how to use the metacognitive strategies when reading correctly. From our point of view, all of these components are required to ensure that participants understand the metacognitive strategies and become aware how to successfully use the strategies when reading belief-relevant multiple texts. Yet, future research might investigate in more depth which of these components might be especially important. This could also provide hints for optimizing the metacognitive strategy training.

5. Conclusion

The digital society and the prominent role of the Word Wide Web as a means for finding information results in new and challenging tasks for readers such as selecting the most appropriate and reliable sources, comprehending a multitude of complex and often inconsistent texts, and making knowledgeable decisions without necessarily possessing sufficient background knowledge. Earlier research and results from the two experiments presented here suggest that prior beliefs play a prominent role during multiple text comprehension, often leading to a text-belief consistency effect. To overcome the text-belief consistency effect and to assist readers in making knowledgeable decisions, including an understanding of belief-inconsistent information is important. The present studies showed that an extended metacognitive strategy training focusing on validation processes during situation model construction and including knowledge provision, observational learning, evaluative feedback, and individual practice, prepares readers for the challenges of comprehending belief-relevant multiple texts.

Acknowledgments

The research reported in this article was supported by the German Research Association (Deutsche Forschungsgemeinschaft, DFG, Grants RI 1100/5-2 and RI 1100/5-3). We would like to thank Caroline Ahrendts and Saskia Kramer for their help in preparing the experiments and collecting the data. The trainings, experimental texts, test items, and data files for the full analyses are available from the authors upon request.

References

Abendroth, J. & Richter, T. (2020a). Text-belief consistency effects in adolescents' comprehension of multiple documents from the Web [El efecto de consistencia en la comprehension lectora de los adolescentes de documentas multiples provenientes de Internet]. *Journal for the Study of Education and Development/ Infancia y Aprendizaje*, *43*, 60-100. doi:10.1080/02103702.2019.1692289

Abendroth, J. & Richter, T. (2020b). Mere plausibility enhances comprehension: The role of plausibility in comprehending an unfamiliar scientific debate. *Manuscript submitted for publication.*

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. doi: 10.1007/s11251-013-9263-8

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. doi: 10.1037/0278-7393.19.5.1061

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. doi: 10.1016/S0022-5371(78)90485-1

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

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). Mahwah, NJ: Lawrence Erlbaum Associates.

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review, 30,* 973-999. doi:10.1007/s10648-018-9436-8

Beker, K., Jolles, D., Lorch, R. F., & van den Broek, P. (2016). Learning from texts: Activation of information from previous texts during reading. *Reading and Writing*, *29*, 1161-1178. doi:10.1007/s11145-016-9630-3

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben [Need for cognition: A scale for assessing commitment and joy regarding cognitive tasks]. *Zeitschrift für Sozialpsychologie, 25,* 147–154.

Bohn-Gettler, C. M., & McCrudden, M. T. (2018). Effects of task relevance instructions and topic beliefs on reading processes and memory. *Discourse Processes, 55*, 410-431. doi:10.1080/0163853X.2017.1292824

Braasch, J. L. G., & Bråten, I. (2017). The Discrepancy-induced Source Comprehension (D-ISC) Model: Basic assumptions and preliminary evidence. *Educational Psychologist, 52*, 167-181. doi: 10.1080/00461520.2017.1323219

Bråten, I., Braasch, J. L., & Salmerón, L. (in press). Reading multiple and non-traditional texts: New opportunities and new challenges. In E. B. Moje, P. Afflerbach, P. Enciso, & N. K. Lesaux (Eds.), *Handbook of Reading Research* (Vol. V). New York, NY: Routledge.

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. doi:10.1080/00461520.2014.916217

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). Mahwah, NJ: Erlbaum.

Cacioppo, J.T., & Petty, R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116-131. doi:10.1037/0022-3514.42.1.116

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1–49. doi:10.3102/00346543063001001

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading*, *42*, 288-325. doi: 10.1111/1467-9817.12269

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Erlbaum.

Dai, D. Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology, 32*, 332 – 347. doi:10.1016/j.cedpsych.2006.05.002

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.

Gilead, M., Sela, M., & Maril, A. (2019). That's my truth: Evidence for involuntary opinion confirmation. *Social Psychological and Personality Science*, *10*, 393–401. doi:10.1177/1948550618762300

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*, 81–112. doi:10.3102/003465430298487

Inquisit 4 [Computer software]. (2015). Retrieved from https://www.millisecond.com.

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 misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1420–1436. doi:10.1037/0278-7393.20.6.1420

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.

Kardash, C. M., & Scholes, R. J. (1996). Effects of preexisting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology, 88*, 260–271. doi:10.1037/0022-0663.88.2.260

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163-182. doi:10.1037/0033-295X.95.2.163

Ku, K. Y., & Ho, I. T. (2010). Metacognitive strategies that enhance critical thinking. *Metacognition and Learning*, *5*, 251-267. doi:10.1007/s11409-010-9060-6

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. doi: 10.1080/07370008.2013.769997

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. doi: 10.1007/s11409-013-9111-x

Maier, J., & Richter, T. (2016). Effects of text-belief consistency and reading task on the strategic validation of multiple texts. *European Journal of Psychology of Education, 31*, 479–497. doi:10.1007/s10212-015-0270-9

Maier, J., Richter, T., & Britt, M. A. (2018). Cognitive processes underlying the text-belief consistency effect: An eye-movement study. *Applied Cognitive Psychology, 32,* 171-185. doi: 10.1002/acp.3391

McCrudden, M. T., & Barnes, A. (2016). Differences in student reasoning about belief-relevant arguments: A mixed methods study. *Metacognition and Learning*, *11*, 275-303. doi:10.1007/s11409-015-9148-0

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.

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation, Vol. 51: The psychology of learning and motivation* (pp. 297-384). San Diego, CA, US: Elsevier Academic Press.

Münchow, H., Richter, T., & Schmid, S. (in press). What does it take to deal with academic literature? Epistemic components of scientific literacy. In O. Zlatkin-Troitschanskaia, H. A. Pant, M.

Toepper & C. Lautenbach (Eds.). *Student learning in German higher education: Innovative modelling and measurement approaches and research results*. Wiesbaden, Germany: Springer.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175–220. doi:10.1037/1089-2680.2.2.175

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, NY: Oxford University Press.

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). Mahwah, NJ: Erlbaum.

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 flexibility* (pp. 125-140). Berlin, Germany: Springer.

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes*, *52*, 337–355. doi: 10.1080/0163853X.2015.1025665

Richter, T. & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A Two-step Model of Validation. *Educational Psychologist, 52*, 148-166. doi:10.1080/00461520.2017.1322968

Richter, T., Münchow, H., & Abendroth, J. (in press). The role of validation in integrating multiple perspectives. In P. Van Meter, A. List, D. Lombardi, & P. Kendeou (Eds.). *Handbook of learning from multiple representations and perspectives*. New York, NY: Routledge.

Richter, T., & Schmid, S. (2010). Epistemological beliefs and epistemic strategies in self-regulated learning. *Metacognition and Learning, 5*, 47-65. doi:10.1007/s11409-009-9038-4

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. doi:10.1037/a0014038

Richter, T., & Singer, M. (2017). Discourse updating: Acquiring and revising knowledge through discourse. In D. Rapp, A. Britt, & M. Schober (Eds.), *Handbook of discourse processes (2nd ed.)* (pp. 167-190). New York: Taylor & Francis.

Rouet, J.-F., Britt, M. A. & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist, 52*, 200-215. doi: 10.1080/00461520.2017.1329015

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19 - 52). Greenwich, CT: Information Age Publishing.

Rouet, J. F., Britt, M. A., Mason, R. A., & Perfetti, C. A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology, 88*, 478–493. doi:10.1037/0022-0663.88.3.478

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. doi:10.1037/0022-3514.32.5.880

Scharrer, L., Stadtler, M. & Bromme, R. (2014). You'd better ask an expert: Mitigating the comprehensibility effect on laypeople's decisions about science-based knowledge claims. *Applied Cognitive Psychology*, *28*, 465-471. doi:10.1002/acp.3018

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. doi:10.1016/0749-596X(86)90002-1

Slavin, R. E. (2014). *Educational psychology: Theory and practice*. Harlow, UK: Pearson.

Stadtler, M., Bromme, R., & Rouet, J.-F. (2018). Learning from multiple documents: How can we foster multiple document literacy skills in a sustainable way? In E. Manalo, Y. Uesaka, & C. Chinn (Eds.), *Promoting spontaneous use of learning and reasoning strategies: Theory, research, and practice* (pp. 46–61). Singapore: Routledge.

Swart, E. K., Nielen T. M. J., Sikkema - de Jong, M. T. (2019). Supporting learning from text: A meta-analysis on the timing and content of effective feedback, *Educational Research Review, 28*, 100296. doi:10.1016/j.edurev.2019.100296

Thomas, E. L., & Robinson, H. A. (1972). *Improving reading in every class: A sourcebook for teachers*. Boston, MA: Houghton Mifflin.

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. doi:10.1016/j.jml.2005.02.001

Wiley, J., Jaeger, A. J., & Griffin, T. D. (2018). Effects of instructional conditions on comprehension from multiple sources in history and science. In In J. Braasch, I. Bråten & M. McCrudden (Eds.), *Handbook of multiple source use* (pp. 341-361). New York, NY: Routledge.

Wolfe, C. R., Britt, M. A., & Butler, J. A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication, 26*, 183–209. doi:10.1177/0741088309333019

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 Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22*, 155-174. doi:10.1007/s10648-010-9134-7

Voss, J. F., Fincher-Kiefer, R., Wiley, J., & Silfies, L. N. (1993). On the processing of arguments. *Argumentation, 7,* 165–181. doi:10.1007/BF00710663

Table 1:

Examples of Paraphrases, Inferences, and Distracters per Topic used in the Verification Task (translated into English)

Type of sentence	Vaccinations	Intelligence
Original paragraph	The fact that many studies are funded by the pharmaceutical industry harms the objectivity of the corresponding research. Researchers commonly need money in order to fund their studies. Thereby, pharmaceutical industries are readily willing to take care of the required funding. The respective studies often lack quality and validity due to the actual goal of the studies: Marketing. *Studies funded by pharmaceutical firms often show more positive results for the respective vaccine than studies that were not funded to such an extent.* Additionally, the vicinity of researchers and the pharmaceutical industry is not an isolated case. […]	Fluid intelligence is effectively trainable. It represents the ability to acquire new knowledge and can be thought of as a tool that is used to build the foundation of knowledge and proper behaviour. Nowadays, challenges that are complex and interconnected require a high level of fluid intelligence in order to cope with them. *For instance, fluid intelligence is needed in order to handle a technical device that has been unfamiliar*. Hence, an effective training of fluid intelligence is required to improve the ability to adapt to new situations and problems that are independent from the training situation and tasks used in the training. […]

Paraphrase	The results of studies funded by the pharmaceutical industry often support the examined vaccine to a higher degree than other studies do.	Fluid intelligence is helpful regarding the operation of unknown technical devices.
Inference	Vaccination studies should no longer be funded by pharmaceutical firms in order to maximize their objectivity.	A training of fluid intelligence positively impacts the practical daily routine.
Distracter	The causal relation between the rubella vaccination and arthritis is empirically validated.	About half the German population has an intelligence quotient between 90 and 110.

Note. The paraphrases refer to the italicized sentences in the original paragraphs. In the original texts, the sentences were not italicized

Table 2:

Means, Standard Deviations and Intercorrelations of Independent Variables (Varied Between-Subjects) and Dependent Variables in Experiment 1

Measure	***M***	***SD***	**1**	**2**	**3**	**4**
1 Training condition (contrast-coded, -1 = PQ4R vs. 1 = metacognitive)	-0.03	1.01	1			
2 Reading order (contrast-coded, -1 = belief-inconsistent first, 1 = belief-consistent first)	0.03	1.01	-.03	1		
3 Situation model strength (belief-consistent text)	1.48	0.78	-.43**	-.15	1	
4 Situation model strength (belief-inconsistent text)	1.37	0.65	-.08	-.26	.44**	1

Note. $N = 37$. Situation model strength: biased-corrected proportion of yes-responses to inference items.

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

Table 3.

Wording of the Explanations on the Three Metacognitive Strategies from Experiment 2 (translated into English).

Metacognitive strategy	Explanation
Awareness of prior beliefs	The first metacognitive strategy refers to the influence of one's prior beliefs on the learning process. It is very common that people have certain beliefs concerning the topic of the texts they are about to read. For instance, everyone probably has an opinion about whether or not cancer check-ups are harmful. Although having opinions and beliefs about controversial topics is good, research has shown that prior beliefs unconsciously influence the evaluation of controversial information concerning their plausibility. Thereby, we tend to reject information that is inconsistent with our beliefs without sufficiently questioning their credibility. This can lead to plausible and well-founded arguments being mistakenly rejected simply because they are inconsistent with one's prior beliefs. Thus, you should become aware of your own beliefs by deliberately questioning them before reading the texts (e.g. 'What is my opinion concerning this topic?'). As soon as you are aware of your beliefs regarding the topic it will be easier for you not to rashly reject belief-inconsistent information. It is rather possible for you to critically question all the information – those consistent and those inconsistent with your prior beliefs.
Monitoring for intertextual	The second metacognitive strategy refers to the monitoring for relationships between the arguments stated in different texts. It is very common that people only consider arguments that are stated within one text disregarding arguments

inconsistencies and relationships	and perspectives of other texts. We tend to read and process new information without connecting them to already established information. Thus, we often fail to perceive textual relationships between arguments of different texts. On the one hand, you should pay attention to paragraphs that are inconsistent between the different texts when identifying relationships between texts. On the other hand, you should focus on textual information that is consistent between the texts. That way you are be able to identify a textual relationship between arguments addressing the pros and cons of cancer check-up you might not have recognized if you would have focused on each argument separately: For instance, the opponents of cancer check-up argue that such examinations can cause severe damages. The proponents do not deny this fact but rather emphasize the great rarity of such consequences. Considering both arguments, it can be inferred that both texts agree upon the possibility of severe consequences resulting from cancer check-ups. The striking difference between the arguments is the stated frequency of such consequences. Identifying relationships and inconsistencies between arguments not only fosters a critical analysis of arguments from the point of view of the arguments in the other text. It also promotes a global comprehension of the scientific controversy addressed in the texts. Hence, the identification of textual relationships is especially important when reading texts that represent different stances concerning a controversial topic.
Active use of prior knowledge	The third metacognitive strategy refers to the active use of one's prior knowledge. It is very common that people possess a certain degree of prior knowledge before reading about a specific topic. For instance, everyone has probably

heard of cancer check-ups before and knows facts about this topic in varying degrees. All facts and figures, that are available in one's memory before reading a certain text, are characterized as one's prior knowledge. Unfortunately, we often experience difficulties in becoming aware of our prior knowledge regarding the topic of the text we are reading. In addition to recalling you own prior knowledge you should actively use those facts and figures in order to identify possible limitations of the presented statements and arguments. Moreover, using your prior knowledge can help you balance conflicting arguments of a scientific controversy and support each stance with profound reasons. In this regard, it is very important for you to explicitly distinguish between proper knowledge and your own personal beliefs concerning the topic of the texts. Since there is a fine line between knowledge and beliefs you should always question whether you are considering actual facts or you own opinion without being able to justify it. Contrary to personal beliefs, your prior knowledge represents a reliable source for evaluating the plausibility of new information. Thus, it is possible for you to critically question arguments and new information within one text as well as between different texts. Therefore, you should become aware of your knowledge concerning the topic of the text and deliberately ask yourself whether it is consistent with the respective argument or not.

Table 4:

Characteristics of the Texts about the Possibility to train Intelligence in Experiment 2.

	Length[a]	Readability[b]	Understandability [c]		Plausibility [c]		Number of arguments [c]		Interest [c]		Clarity of Stance [c]	
			M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Pro Text	828	38	4.52	0.60	4.14	0.69	3.50	0.97	4.20	0.79	4.90	1.60
Contra Text	867	38	4.86	0.40	4.42	0.51	3.55	1.21	4.30	0.82	5.60	0.70

N = 10.

Note. Understandability (Cronbach's α = .79/.83) and plausibility were measured with nine items each (Cronbach's α = .85/.87). Number of Arguments was indicated by the number of identified arguments in an open answer question. Interest and clarity of stance were assessed with one item each. All response categories ranged from 1 (*completely disagree*) to 6 (*completely agree*).

[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 from ten university students.

Table 5:

Means, Standard Deviations and Intercorrelations of Independent Variables (Varied Between-Subjects) and Dependent Variables in Experiment 2

Measure	***M***	***SD***	**1**	**2**	**3**	**4**	**5**	**6**
1 Training condition (contrast-coded, -1 = PQ4R vs. 1 = metacognitive)	.09	1.00	1					
2 Reading order (contrast-coded, -1 = belief-inconsistent first, 1 = belief-consistent first)	.04	1.01	.08	1				
3 Text Topic (contrast-coded, -1 = intelligence, 1 = vaccination)	0.00	1.01	-.04	.09	1			
4 Need for cognition [a]	5.02	0.78	-.06	-.09	-.17	1		
5 Situation model strength (belief-consistent text)	2.22	0.83	-.11	.13	-.12	.47**	1	
6 Situation model strength (belief-inconsistent text)	1.88	0.78	.16	.08	.05	.32*	.43**	1

Note. $N = 46$. Situation model strength: biased-corrected proportion of yes-responses to inference items. [a] *M* and *SD* are provided for the raw need for cognition scores.

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

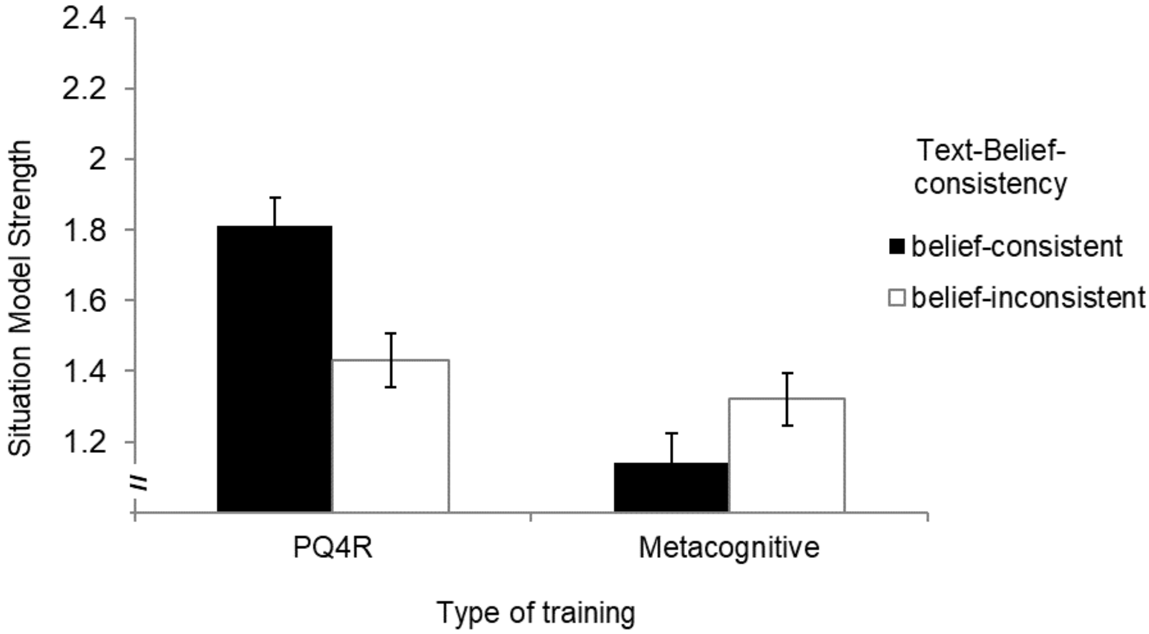

Figure 1. Effects of text-belief consistency and training condition on situation model strength in Experiment 1. Error bars represent the standard error of the mean.

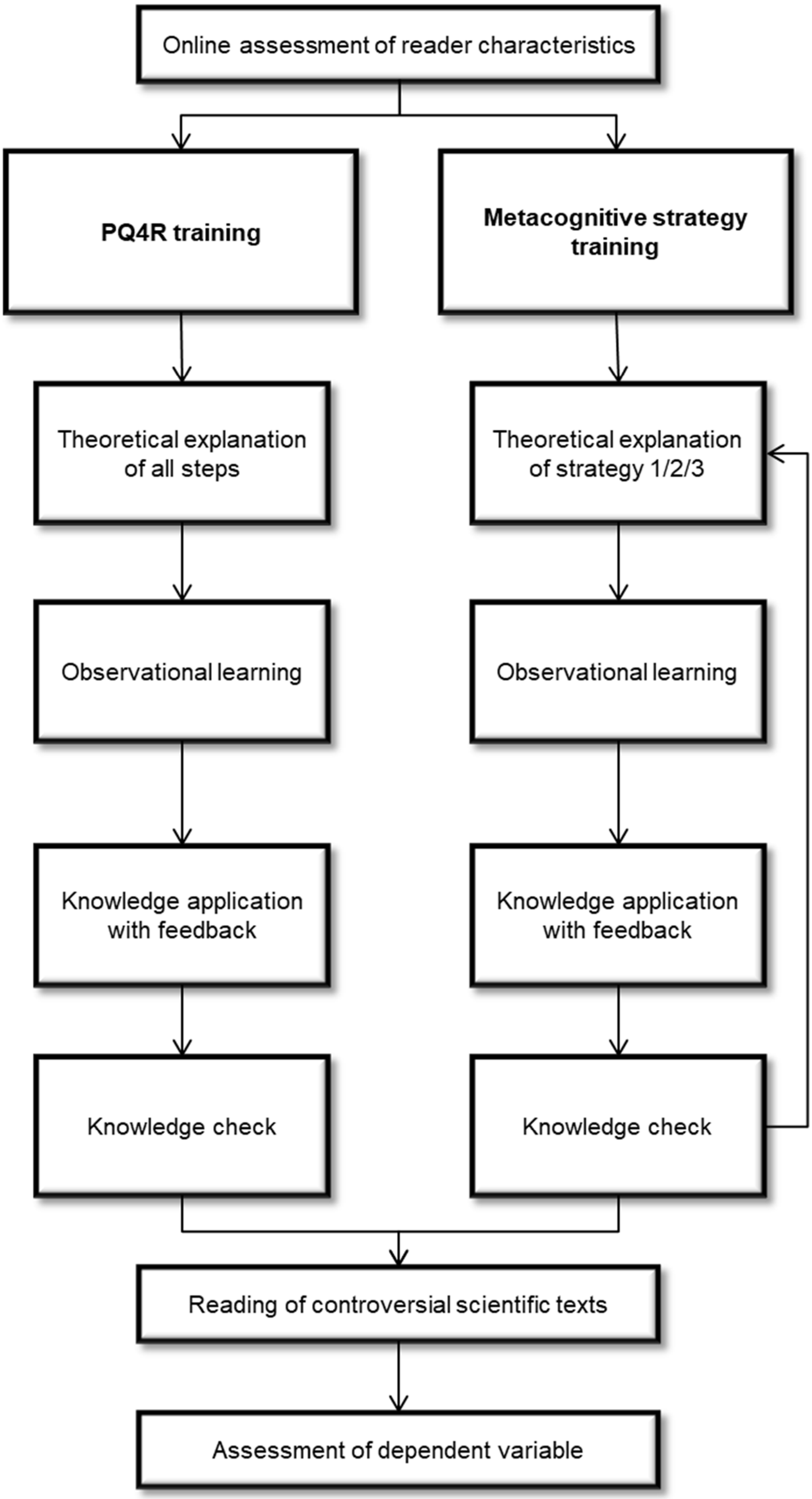

Figure 2. Procedure of Experiment 2.

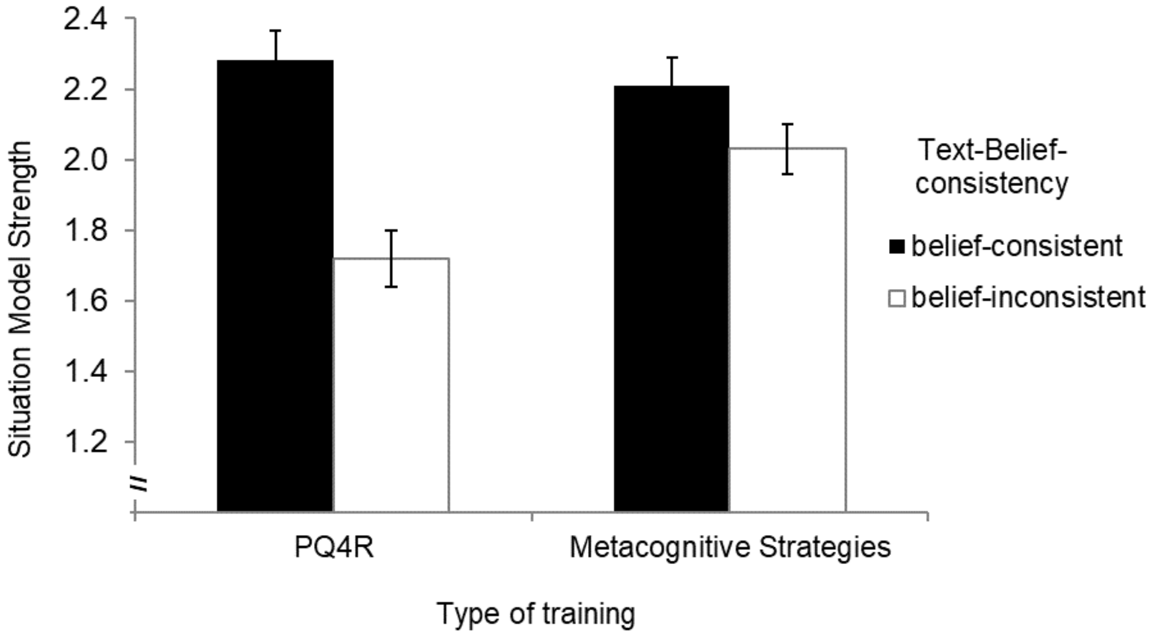

Figure 3. Effects of text-belief consistency and training condition on situation model strength in Experiment 2. Error bars represent the standard error of the mean.