2018

Effects of Task Relevance Instructions and Topic Beliefs on Reading Processes and Memory

Catherine Bohn-Gettler
College of Saint Benedict/Saint John's University, cbohngettler@csbsju.edu

Matthew McCrudden

Recommended Citation

Bohn-Gettler, Catherine M.; McCrudden, Matthew T. (2018). Effects of Task Relevance Instructions and Topic Beliefs on Reading Processes and Memory *Discourse Processes, 55*(4), 410-431. http://dx.doi.org/10.1080/0163853X.2017.1292824

The Effects of Task Relevance Instructions and Topic Beliefs on Reading Processes and Memory

Catherine M. Bohn-Gettler

College of Saint Benedict – Saint John's University

Matthew T. McCrudden

Victoria University of Wellington

Author Note

Catherine M. Bohn-Gettler, Education Department, College of Saint Benedict – St. John's University; Matthew T. McCrudden, School of Education, Victoria University of Wellington

We thank Brady Marzolf, Melinda Mueller, Ariel Glover, and Dinithi Perera for their assistance in conducting this study.

Correspondence concerning this article should be addressed to Catherine M. Bohn-Gettler, Education Department, College of Saint Benedict – St. John's University, 37 South College Avenue, Saint Joseph, MN 56374. Email: cbohngettler@csbsju.edu; Phone: (320) 363-5087.

The Effects of Task Relevance Instructions and Topic Beliefs on Reading Processes and Memory

Abstract

This study investigated the effects of task relevance instructions and topic beliefs on reading processes and memory for belief-related text. Undergraduates received task instructions (focus on arguments for versus against) before reading a dual-position text. In Experiment 1 (*n* = 88), a reading time methodology showed no differences in reading time for task-relevant and task-irrelevant text, but participants recalled task-relevant text better than task-irrelevant text independently of whether the information was consistent with their topic beliefs. In Experiment 2 (*n* = 76), a think-aloud methodology showed that participants engaged in confirmation strategies when reading belief-consistent text and disconfirmation strategies when reading belief-inconsistent text, independently of whether the information was relevant to their task instructions. Nonetheless, participants recalled task-relevant text better than task-irrelevant text. The results indicate that task relevance instructions affect memory independently of beliefs, but that beliefs affect processing independently of task relevance instructions. Thus, moment-by-moment reading processes and memory for text can operate differently as a function of topic beliefs.

Keywords: reading processes, beliefs, task relevance instructions, think-aloud, memory

The Effects of Task Relevance Instructions and Topic Beliefs on Reading Processes and Memory

Students have the right to believe whatever they want to believe, and schools cannot require students to change their beliefs (Moshman, 2009). However, students must take courses, deemed necessary by appropriate authorities on the basis of academic justifications, to complete particular qualifications (e.g., high school diploma). This can sometimes create a tension between what students believe and what they are expected to know. For instance, some students believe that the diversity of life on Earth can be explained on the basis of religion. However, science teachers expect students to know evidence and arguments for biologists' position that evolution explains genetic changes that occur in a population over time.

It is important for students to understand content, independently of their beliefs. Achieving this aim can be difficult because topic beliefs can affect how people process belief-related information in a variety of ways (Sinatra, Kienhues, & Hofer, 2014; Wolfe, Tanner, & Taylor, 2013). For instance, individuals may primarily focus on information that reinforces their beliefs (e.g., Hart et al., 2009), primarily focus on information that challenges their beliefs in an attempt to disconfirm it (Eagly & Chaiken, 1993; Edwards & Smith, 1996; Wolfe et al., 2013), or focus on both information that reinforces and challenges their beliefs (Taber & Lodge, 2006).

Given that reading is a primary means of school-based learning, and that reading processes are related to memory (e.g., Goldman & Varma, 1995; Kintsch, 1998; McNamara & Magliano, 2009), it is important to identify ways to promote students' understanding of belief-related information. As one example, if students only focus on information that reinforces beliefs, they may fail to understand information that challenges their beliefs. Therefore, it is important to investigate ways to promote student understanding of belief-related text. The purpose of this study was to investigate the effects of task relevance instructions and topic beliefs on moment-by-moment reading processes and memory for belief-related text.

Goal Focusing Model

Task relevance instructions orient readers to an assigned reading task, such as when a teacher asks students to read a text for a particular purpose (Gil, Bråten, Vidal-Abarca, & Strømsø, 2010; Vidal-Abarca, Mañá, & Gil, 2010). For example, a teacher may ask students to read a dual-position text to help them understand two sides of a controversial issue. In such a situation, the teacher may give students additional instructions to help direct their attention, which can affect how they process and recall the text (McCrudden, Magliano, & Schraw, 2010; Rouet & Britt, 2011; van den Broek, Bohn-Gettler, Kendeou, Carlson, & White, 2011; van den Broek, Lorch, Linderholm, & Gustafson, 2001).

The goal-focusing model of relevance describes the events that take place when students are given task instructions for an assigned reading task (McCrudden, Magliano, & Schraw, 2011; McCrudden & Schraw, 2007). When a reading task is assigned, relevance cues and reader intentions can affect online reading processes and offline products (i.e., changes in memory or learning that result from reading). Relevance cues are signals communicated by an external source (e.g., teacher) that indicate the extent to which information is relevant to a task, such as instructions to focus on specific types of information. These cues are meant to help readers determine how to process the text while they read and how they may use information after they read. However, reader intentions can affect how readers process and remember text. When given assigned reading tasks, readers bring their knowledge, beliefs, values, expectations, and experiences, which can affect their decisions about what and how to process text information.

Relevance cues and reader intentions jointly affect readers' standards of relevance. *Standards of relevance* are the criteria readers use to determine the relevance, or instrumental value, of ideas in a text in relation to their goals (McCrudden et al., 2010). A segment that has greater perceived value is deemed more relevant, whereas a segment that has less perceived value is deemed less relevant. Moreover, information that a reader perceives to be relevant to a goal may be different from information that is important to the coherence of the text (McCrudden & Schraw, 2007; Schraw, Wade, & Kardash, 1993) or what a teacher, for example, considers to be important (Alexander & Jetton, 2000). Thus, goals could reflect relevance cues, reader intentions, or both. Therefore, standards of relevance may lead one to focus on and process information from a text, and this focus may differ from the purpose for which the text was written, or may differ from the originally assigned task. Readers' goals and standards of relevance affect how they allocate attention and their use of strategic processing during reading, such that they tend to allocate more attention to goal-relevant information and utilize effortful strategic processing activities when reading this information (Kaakinen & Hyönä, 2011; Kaakinen, Hyönä, & Keenan, 2002; McCrudden & Schraw, 2007). Further, the mental model that results from reading tends to reflect information that is deemed more relevant to readers' goals.

Nonetheless, readers who are given the same task instructions and read the same text may develop different goals, process the text differently, and construct different mental representations of the text. For instance, in McCrudden et al. (2010), undergraduates read a text about several countries, and were asked to focus on information about a specific country. Collectively, participants spent more time reading task-relevant information, and they remembered this information better than task-irrelevant information. However, follow-up interviews indicated that participants implemented task instructions differently, such that some students focused exclusively on task-relevant information, whereas others focused on both task-relevant and task-irrelevant information. Reading time and recall data corroborated the interview data. Thus, readers do not uniformly use task instructions for reading.

This underscores the idea that many factors can influence how readers approach texts in educational settings (Bohn-Gettler & Kendeou, 2014; McCrudden & Schraw, 2007; Rapp & van den Broek, 2005; van den Broek, Risden, & Husebye-Hartmann, 1995), and that task instructions do not uniformly affect readers' strategies. Readers can develop different goals and processing strategies in response to the same task instructions (Kendeou, Bohn-Gettler, & Fulton, 2011; McCrudden et al., 2010). One factor, topic beliefs, may play an important role in how readers enact explicit task instructions.

Topic Beliefs

A belief is an idea that a person accepts to be true, but that does not require verification from others (Murphy & Mason, 2006). Beliefs held about particular topics (i.e., topic beliefs) can affect online processes and offline products; however, previous research shows that results differ regarding when and how this may occur. Edwards and Smith (1996) found that participants spent more time reading belief-inconsistent arguments. After reading, they listed more refutational statements for belief-inconsistent arguments, but listed more supportive statements for belief-consistent arguments. Similarly, in Kardash and Howell (2000), participants who thought-aloud while reading a dual-position text tended to refute, disagree with, or make more judgements about belief-inconsistent information than belief-consistent information. Taber and Lodge (2006) found that participants spent more time on belief-consistent information. In contrast, Maier and Richter (2013) found that readers had longer reading times for belief-inconsistent information when texts were presented in block format (i.e., one position presented in its entirety before the other position). Nonetheless, these readers encoded belief-inconsistent and belief-consistent texts differently. Readers more accurately identified paraphrases for belief-inconsistent information, but more accurately identified inferences about belief-consistent information on a sentence recognition task. However, these online and offline differences did not occur when the texts were interleaved. Wolfe et al. (2013) found that argument strength primarily affected reading time, independently of whether the arguments were consistent with participants' beliefs. Thus, topic beliefs can affect online processes and offline products, but there are differences in how and when this occurs.

Nonetheless, previous research indicates that task instructions can affect how readers process content independently of their beliefs. Maier and Richter (2015) asked undergraduates to read belief-related texts for the purpose of summarizing the text or generating an argument. When asked to summarize, participants tended to use memorization strategies for belief-consistent information, and spent more time reading this information. Conversely, when asked to generate an argument, reading times for belief-consistent and belief-inconsistent information did not differ; however, participants made more elaborations when reading belief-inconsistent texts, which coincided with better comprehension of belief-inconsistent information. Thus, task instructions can affect how readers process belief-related content.

Overview of the Present Study

The purpose of this study was to use the three-pronged approach (Magliano & Graesser, 1991) to investigate the effects of task relevance instructions and topic beliefs on reading processes and memory for text. We used theory, processing data (i.e., reading times and verbal protocols), and recall data to investigate online processes and offline products when individuals read belief-related text.

In Experiment 1, participants read the text silently via computer and then recalled the text. Reading time per text segment was recorded. We used a reading time methodology to capture both strategic and automatic reading processes. In Experiment 2, we used a think-aloud methodology to assess the types of processes readers use when they read task-relevant and task-irrelevant information, and participants recalled the text after they read. Although a think-aloud methodology is more intrusive than a reading time methodology, reading time does not reveal the cognitive processes readers use during reading. Thus, reading time and think-aloud data together provided a more comprehensive assessment of reading processes than either methodology alone (Kendeou & van den Broek, 2007; Magliano & Graesser, 1991).

We chose the topic of intelligent design (an argument for an intelligent creator of life on earth, often interpreted as a God) because many people hold beliefs on the topic, and the sentences in the text could be separated into two discrete categories; sentences that support and sentences that oppose teaching ID in science classrooms. When two categories of sentences differ with respect to belief-consistency, it is possible to determine whether task instructions affect how readers process and remember belief-consistent and belief-inconsistent information.

The present study aims to extend previous research in two main ways. First, we applied the three-pronged approach (Magliano & Graesser, 1991) to examine both online processes and offline memory associated with reading an extended belief-related text. To accomplish this, we used both reading time and think-aloud data to more clearly investigate moment-by-moment online processing of an extended text, and recall data to investigate offline products of reading. Previous research has used reading times for single-sentence arguments of unrelated topics

(Edwards & Smith, 1996), which may lead to different types of processing than connected discourse on the same topic. Research has also used post-reading thought listing tasks (Edwards & Smith, 1996), which may differ from moment-by-moment processing. Think-aloud protocols allow readers to articulate thoughts at the time in which they occur rather than retrospectively after reading. This is important for measuring thoughts for extended discourse, because memory for specific thoughts across an entire text may be more accessible during reading and less subject to memory difficulties that may occur after reading. The reading time methodology enabled us to assess time spent reading different categories of text information and rule out differences in memory as a result of time spent reading. The think-aloud methodology enabled us to assess overt cognitive processes that may have occurred and corresponded to reading time data. That is, the reading time and think-aloud methodologies enabled us to investigate strategic and automatic reading processes differently. Experiment 1 examined reading times and recall, whereas Experiment 2 examined think-alouds and recall.

Second, we used task instructions to target specific categories of information. Conversely, Maier and Richter (2015) used general task instructions that asked participants to read for a particular purpose (i.e., read to summarize or read to form an argument). General instructions allow readers to interpret general, less-specific task instructions with greater variability. However, the current study utilized specific instructions that focused readers on particular categories of information. Readers may respond differently to specific task instructions because specific instructions provide more explicit criteria for distinguishing between task-relevant and task-irrelevant information.

Experiment 1

The purpose of Experiment 1 was to investigate the effects of topic beliefs and task instructions on reading time and memory. Topic beliefs were used to determine belief/task-alignment; that is, the combination of whether the readers' task instructions were to focus on belief-consistent information (if the task instructions targeted belief-consistent information, there was a belief/task-match), belief-inconsistent information (if the task instructions targeted information that was inconsistent with a reader's beliefs, there was a belief/task-mismatch), or belief-neutral information (if the reader was ambivalent about the topic, the instructions were belief/task-neutral).

On the one hand, task instructions might affect readers' online processing and offline products independently of their beliefs. The task instructions could give readers criteria for determining the extent to which text segments are relevant to their task, which may lead to longer reading times for and better recall of task-relevant information compared to task-irrelevant information (Kaakinen & Hyönä, 2005, 2011; Kaakinen et al., 2002; Kaakinen, Hyönä, & Keenan, 2003; McCrudden & Schraw, 2007, 2010). If this is the case, participants in the belief/task-match group should spend more time on belief-consistent information and remember this information better than belief-inconsistent information. And, participants in the belief/task-mismatch group should spend more time on belief-inconsistent information and remember this information better than belief-consistent information.

On the other hand, task instructions might not affect readers' online processing and offline products. Rather, beliefs may exert a more powerful influence than task instructions, and could affect processing and memory in at least two ways. First, individuals might attempt to bolster their beliefs through selective exposure to belief-consistent information by attempting to confirm the information (Hart et al., 2009). If this is the case, participants should have longer reading times and better memory for belief-consistent than belief-inconsistent text, independently of whether the information is task-relevant. Alternatively, individuals might attempt to bolster their beliefs through selective exposure to belief-inconsistent information by attempting to disconfirm the information (Edwards & Smith, 1996). If this is the case, participants should have longer reading times and better memory for belief-inconsistent than belief-consistent text, independently of whether the information is task-relevant.

Method

Participants and context. Participants were 93 undergraduates (50.5% female) at a medium-sized public university in Kansas, located in the Midwestern part of the United States. The mean age in years was 23.3 (*SD* = 8.3). Participants were recruited from psychology and education classes, and received course credit for their involvement in the study.

Debate about the teaching of evolution and creationism (e.g., intelligent design) has a history in Kansas. For instance, in 2005, the Kansas State Board of Education (KSBE) approved a draft of science curriculum standards requiring evolutionary theory and intelligent design be taught for equal amounts of time in science classrooms (although this draft was eventually rejected). As recently as 2013, an anti-evolution group sued the KSBE for including science curriculum that teaches evolution on the ground that excluding ID from science classrooms discriminates against religion. Thus, this topic was relevant to the general context in which the study was conducted.

Design and conditions. Figure 1 displays the study design. Participants were randomly assigned to task instructions, to either focus on arguments *for* or *against* teaching ID in science classrooms. We also collected a measure of participants' pre-existing beliefs about whether ID should be taught in science classrooms (and beliefs were categorized as *for, against,* or *neutral* toward the topic). In this way, we were able determine whether participants' beliefs aligned or did not align with their task instructions (i.e., belief/task-alignment). Hence, belief consistency as a variable was measured by the combination of task instructions and pre-existing topic beliefs. For instance, if a participant believed ID *should* be taught, and the task instructions targeted arguments *for* teaching ID, *for* arguments were relevant and consistent with participants' beliefs, resulting in a match between beliefs and task instructions. However, if a participant believed ID *should* be taught, and the task instructions targeted arguments *against* teaching ID, *against* arguments were relevant but inconsistent with participants' beliefs, resulting in a mismatch between beliefs and task instructions. If a person was *neutral* toward the topic of teaching ID, the belief/task-alignment was neutral (regardless of the task instructions).

We also determined whether each text segment was relevant or irrelevant to the task instructions. For instance, if the task instructions targeted arguments *for* teaching ID, text segments arguing *for* teaching ID were task-relevant, whereas text segments arguing *against* teaching ID were task-irrelevant. Taken together, we used a 3 (belief/task-alignment: match, mismatch, vs. neutral; between subjects) x 2 (text segment type: task-relevant vs. task-irrelevant; within subject) mixed model design.

Participants were randomly assigned to a task instruction condition before beliefs were measured; therefore, there was an uneven distribution of participants for the different levels of belief/task-alignment. Table 1 provides the number of participants as a function of task instructions and beliefs. Also, three participants' data were removed from the study, and two participants did not answer the question assessing their beliefs (described later). This design produced three conditions: (a) belief/task-match (n = 25), (b) belief/task-mismatch (n = 33), (c) belief/task-neutral (n = 30).

Materials.

Topic beliefs. The topic beliefs instrument measured participants' beliefs about whether intelligent design should be taught in science classrooms. Participants read the following background information: "Intelligent design (ID) is an argument for the existence of a creator for life on Earth, often interpreted as a God. According to the Discovery Institute, ID is the view that 'certain features of the universe and of living things are best explained by an intelligent cause, not an undirected process such as natural selection.'" Next they rated their agreement with the following statement, "I think intelligent design should be taught in science classrooms" on a 9-point Likert-type scale (1 = strongly agree; 9 = strongly disagree). Responses were used to determine whether participants were more-accepting of (n = 29; ratings of 1 through 3), neutral/ambivalent towards (n = 30; ratings of 4 through 6), or less-accepting of (n = 29; ratings of 7 through 9) ID being taught in science classrooms. Two participants did not answer the topic belief question.

Text. The text described arguments for and against teaching ID in science classrooms (1427 words, Flesch-Kincaid grade level 12.2) and was a composite of various arguments from articles and blogs on the Internet. The introduction (8 sentences, 158 words) began by introducing the topic and indicated that some people think ID should be taught in science classes, whereas others do not, and that the text included arguments for both sides of the issue.

Each argument began with a claim (e.g., some people are in support of/opposition to teaching ID) and a reason (e.g., ID is not an empirical scientific theory, and hence does not belong in science classrooms), followed by a detailed explanation of the reason. The four arguments for teaching ID consisted of 26 sentences (620 words), and the four arguments against teaching ID consisted of 27 sentences (626 words). An argument in favor of teaching ID was followed by an argument against it, although successive arguments were not necessarily related (i.e., the text did not follow an argument-counterargument structure). This was done to minimize text-belief consistency bias, because integrated (i.e., interleaved) formats reduce belief-consistency effects in comparison to blocked formats (Maier & Richter, 2013; Wiley, 2005). The text concluded with a one-sentence statement (23 words) indicating that the issue would likely continue to be controversial.

Apparatus. The experiment was conducted utilizing a Dell desktop computer with E-Prime software (E-Prime, Psychological Software Tools, Inc.). Participants sat at a desk in front of a color monitor, and rested their right hand on the mouse. The text was presented in standard lower- and upper-case type, and was centered on the screen.

Procedure. Each participant completed the session independently. Participants first completed the topic belief instrument. Then, they read a practice text, in which phrases were presented one at a time, on a computer screen. A phrase was defined as a sentence or partial sentence that contained a subject and verb, and that communicated one idea (i.e., Kendeou, Bohn-Gettler, White, & van den Broek, 2008). Participants proceeded from one phrase to the next, at their own pace, by pressing the mouse button with their right index finger. They were not able to re-read prior text. Reading times in milliseconds (ms) were collected.

Next, participants were randomly assigned a task instruction condition. The two task instruction conditions were to either to focus on arguments *for* or focus on arguments *against* teaching ID in science classrooms. All participants were informed that they would read a text about reasons for and against teaching ID in science classrooms and then complete a comprehension task after reading.

After they read, participants were asked to verbally retell the text as if they were speaking to a fellow student who had not read the text. No other instructions were provided for the retelling in an effort to minimize demand characteristics. The experimenter provided no prompts other than to ask at the end of the recall, "Would you like to add anything else?" Responses were audio recorded and transcribed.

Screening of reading time data. Reading times less than 10ms, or greater than 3 SD above the mean were removed, resulting in removal of 1.6% of the data. Descriptive analyses identified two participants as outliers whose data were removed from the study. One participant's overall reading time was 3 SD above the mean, and one's overall reading times was 2 SD below the mean. Another participant's data was removed due to self-reported vision issues and loud construction noises that interfered with reading. Finally, two participants did not answer the topic belief question.

Coding of recall data. Three researchers parsed the participants' recall transcripts into idea units, generally defined as subject-verb phrases that communicated one idea (Kendeou et al., 2008). Because a sentence could convey multiple ideas, several idea units could occur within a sentence. Hence, this provided a more accurate measure of the number of ideas from a text the participant included in their recall. The participants' idea units were matched to the corresponding idea within the text (the text contained 165 possible ideas: 21 introduction/conclusion, 71 related to pro arguments, and 73 related to against arguments). If the participant included an idea unit more than once, it was only counted once to ensure that only unique ideas were included. This gave us a better indication of how many unique ideas from each respective position that participants recalled. Although we scored for the entirety of the text, we only analyzed the percentage of unique ideas recalled for the eight arguments, and not the neutral introductory or conclusion content. Hence, we broke down the percentage recalled by whether the text was relevant or irrelevant to their task instructions. The experimental conditions of the participants were masked to the raters during the coding process. The percentage agreement among raters was 91% (Fleiss' *kappa* = .79).

Results

We conducted a separate 3 x 2 mixed model univariate analyses of variance (ANOVA) with belief/task-alignment (which represented the combination of task instructions and participants' pre-existing beliefs: match, mismatch, or neutral) as a between-subjects variable, and segment type (task-relevant or task-irrelevant) as a within-subject variable on reading time and recall. ANOVAs are fairly robust for accounting for unequal sample sizes, but we report the Greenhouse-Geisser statistics to account for this. We computed partial eta squared (η^2) for the measurement of effect size, with partial η^2 qualifying values of approximately 0.01 as small effects, values of 0.06 as medium effects, and values of approximately 0.14 or more as large effects (see Olejnik & Algina, 2000). All follow-up post-hoc tests utilized Tukey's HSD, which controls for family wise error at the .05-level. We only report post-hoc tests that were significant at the $p < .05$ level of significance unless otherwise noted.

Reading times. Table 2 contains descriptive statistics related to the reading times, which are reported as time per word (ms)[1]. We only analyzed reading times for the sentences containing arguments for or against teaching ID in science classrooms (i.e., excluded the

[1] Analyses were also conducted utilizing reading times per syllable, as well as reading times per character. In all instances, the pattern of results was the same as reading times per word. We opted to report reading times per word for ease of interpretation.

introductory paragraph and the conclusion sentence at the end of the entire passage). The interaction was not significant [$F(2, 85) = .52$, $p = .60$, $\eta^2 = .01$], nor were the main effects for belief/task-alignment [$F(2, 85) = 2.24$, $p = .11$, $\eta^2 = .05$] or text segment [$F(1, 85) = 1.18$, $p = .28$, $\eta^2 = .01$].

Recall. Table 3 contains descriptive statistics related to the recall data, which are reported as proportion of unique text ideas recalled. There were 165 possible ideas in the text; participants recalled approximately 7-9% of the text (11 to 15 unique ideas). The interaction was not significant, $F(2, 85) = 1.61$, $p = .20$, $\eta^2 = .04$. The main effect for segment type was significant, $F(1, 85) = 35.03$, $p < .001$, $\eta^2 = .29$. Participants recalled more task-relevant segments than task-irrelevant segments. For instance, participants in the belief/task-match group recalled more task-relevant information (which was belief-consistent) than task-irrelevant information (which was belief-*inconsistent*). Similarly, participants in the belief/task-mismatch group also recalled more task-relevant information (which was belief-*inconsistent*) than task-irrelevant information (which was belief-consistent). The main effect for belief/task-alignment was not significant, $F(2, 85) = .90$, $p = .41$, $\eta^2 = .02$.

Discussion

There was no difference in reading time between task-relevant and task-irrelevant text; however, participants recalled task-relevant text better than task-irrelevant text independently of their pre-existing topic beliefs. These findings indicate that task instructions promoted memory for task-relevant information, independently of whether the information conflicted with readers' beliefs. These differences in memory could not be attributed to longer reading times for task-relevant information. One explanation for this outcome is that specific task instructions provide readers with clear criteria for developing standards of relevance during reading and that such criteria facilitated memory for task-relevant information.

These findings also suggest that one way to promote memory of information that differs from one's beliefs is to ask the person to focus specifically on that information. However, similarities in reading times do not necessarily indicate that participants engaged in similar reading processes (Magliano & Graesser, 1991). Prior work has produced mixed effects with regard to the effects of beliefs on reading times. For example, Maier and Richter found that reading times were longer for belief-consistent information when reading a blocked (but not interleaved) text (2013), or when reading with the goal of summarizing, but not when reading to build an argument (2015). Wolfe et al. (2013) found that reading times varied as a function of the strength of arguments, not beliefs. However, the present study found no effects for reading time, despite prior work that relevance instructions encourage readers to spend more time reading task-relevant information (Kaakinen & Hyönä, 2011; Kaakinen et al., 2002; McCrudden & Schraw, 2007). The results from the present study therefore point to a need to understand why neither beliefs nor task instructions affected reading times, because even though reading times were similar for all groups, such similarities could be masking very different processes occurring during reading. It may be that participants applied different processes as a function of task instructions and beliefs, such as evaluations, text-, or knowledge-based inferences while reading belief-consistent and belief-inconsistent viewpoints (Maier & Richter, 2015). Therefore, we conducted a second experiment using a think-aloud methodology to investigate online reading processes.

Experiment 2

The purpose of Experiment 2 was to investigate the effects of task instructions and topic beliefs on online processes using a think-aloud methodology and offline products via free recall. In think-aloud tasks, verbalized thoughts provide direct insights into readers' moment-by-moment cognitive processes (Ericsson & Simon, 1993; Magliano, Trabasso, & Graesser, 1999). While many cognitive processes that occur during reading are automatic, proficient readers are typically aware of explanatory processes (Graesser, Singer, & Trabasso, 1994) and tend to report them when thinking-aloud (Magliano & Millis, 2003; Magliano et al., 1999; Trabasso & Magliano, 1996). Thinking-aloud has received extensive validation as a tool for revealing comprehension processes during reading (Coté & Goldman, 1999; Magliano et al., 1999; Pressley & Afflerbach, 1995). The cognitive processes that can facilitate text comprehension include text rehearsal, making inferential connections between textual information, elaborating on text utilizing prior knowledge, evaluating the quality of text arguments, and more (Ericsson & Simon, 1993; Pressley & Afflerbach, 1995).

We expected that task instructions would lead to better recall for task-relevant information across conditions. This would lend further support for the effectiveness of task instructions on memory. However, it was unclear what cognitive processes led to this outcome in Experiment 1, particularly given that reading times did not differ between task-relevant and task-irrelevant text. There are at least four possible explanations.

One possibility is that participants were motivated to disconfirm belief-inconsistent information, independently of task instructions. If this was the case, refutational evaluations should be used more than supportive evaluations for belief-*inconsistent* text, whereas there should be no difference in their use for belief-*consistent* text. Further, backward inferences and elaborations should be used more at belief-inconsistent text than at belief-consistent text because participants are accessing textual information or prior knowledge to undermine belief-inconsistent content (similar to Edwards & Smith, 1996).

A second possibility is that participants were motivated to confirm belief-consistent information, independently of task instructions. If this was the case, supportive evaluations should be used more than refutational evaluations at belief-*consistent* text, whereas there should be no difference in their use at belief-*inconsistent* text. Further, backward inferences and elaborations should be used more at belief-consistent text than at belief-inconsistent text because participants are accessing textual information or prior knowledge to bolster belief-consistent content (aligning with research involving rating the quality of arguments, such as Edwards & Smith, 1996; Hart et al., 2009; Klaczynski & Robinson, 2000; Taber & Lodge, 2006).

A third possibility is that participants were motivated to both disconfirm belief-inconsistent text and confirm belief-consistent text, independently of task instructions (e.g., Taber & Lodge, 2006). If this is the case, refutational evaluations should be used at belief-inconsistent text, whereas supportive evaluations should be used at belief-consistent text. There should be no differences in the use of bridging inferences or elaborations for belief-inconsistent and belief-consistent text.

A fourth possibility is that participants' engaged in differential think-aloud processes as a function of task instructions, but not beliefs. And, the differential processes embodied themselves in similar reading times. If this is the case, participants may provide supportive evaluations for task-relevant text, and refutational evaluations for task-irrelevant text, independently of their pre-existing beliefs. This possibility seemed most likely for the belief/task-neutral groups given they were generally ambivalent about the topic. For the belief/task-match and belief/task-mismatch groups, any of the previously described possibilities could emerge.

Method

Participants and context. Participants were 80 undergraduates (65% female) from the same medium-sized public university in the US state of Kansas as in Experiment 1. None of the participants from Experiment 1 participated in Experiment 2. The mean age in years was 23.6 (*SD* = 8.01).

Design and conditions. The design and conditions were the same as used in Experiment 1. We used a 3 (belief/task-alignment: *match*, *mismatch*, vs. *neutral*) x 2 (segment type: *relevant* vs. *irrelevant* to task instructions) mixed model design. Table 1 provides the number of participants as a function of task instructions and beliefs. As before, there was an uneven distribution of participants for the different levels of belief/task-alignment. The design produced three conditions: (a) belief/task-match (*n* = 31), (b) belief/task-mismatch (*n* = 23), and (c) belief/task-neutral (*n* = 22).

Materials.

Topic beliefs. The topic beliefs instrument was the same as used in Experiment 1. Responses were used to determine whether participants were more-accepting of (*n* = 20), neutral/ambivalent towards (*n* = 24), or less-accepting of (*n* = 36) ID being taught in science classrooms.

Text. The text was the same text as used in Experiment 1.

Procedure. Each participant completed the session independently. The procedure was similar to Experiment 1 with one main difference: Instead of reading the text silently, the participants thought-aloud about the text. Participants first completed the topic belief instrument and then did a practice think-aloud task. Some sentences were marked by stars **(***)** indicating that the participant was required to think-aloud at that particular sentence; however, they were free to think-aloud at any point. The stars occurred after the last sentence of the introduction, after the first sentence of each argument section (in which the main idea of the argument was presented), and after the final summative sentence of the text (which contained neutral content). Hence, the stars were equally spread across the arguments. Each paragraph was presented on one page, but each sentence was numbered and presented on its own line. The experimenter modelled reading a text aloud, sentence-by-sentence, and verbally stated his or her thoughts after each sentence for the first half of the practice text. The experimenter demonstrated various productions, with examples drawn from a rubric developed to exemplify most possible processes (including elaborations, evaluations, text rehearsal, and backward inferences). For the second half of the practice text, participants practiced reading and thinking aloud. If the participant forgot to think out loud at a starred sentence, the experimenter asked a non-leading question, such as "What are you thinking after you read this sentence?" The experimenter did not answer questions or decode words (Ericsson & Simon, 1993).

Next, participants were randomly assigned either: (a) to focus on arguments *for* (n = 40), or (b) to focus on arguments *against* (n = 40) teaching ID in science classrooms. They received the same instructions as described in Experiment 1. Participants proceeded to think-aloud about the text in the same manner as described for the practice text. After reading, participants recalled the text using the same method as described in Experiment 1, and responses were recorded and transcribed.

Coding of think-aloud data. Participants' think-aloud responses were parsed into idea units and coded by three raters. The experimental conditions of the participants were masked to the raters. The response categories were adapted from Bohn-Gettler and Rapp (2011). This analysis specifically focused on evaluations, backward inferences, and elaborations. We focused on these processes because they can encourage comprehension beyond text rehearsal (van den Broek et al., 2001), have been implicated in previous research on beliefs (Maier & Richter, 2015), and occurred frequently enough in the data to warrant analyses. *Evaluations* occurred when participants stated opinions about the content of the text, and were further coded as supportive, refutational, or neutral toward the text segment. *Backward inferences* occurred when participants referred to information presented in earlier sentences that was relevant or explained the current sentence. *Elaborations* occurred when participants retrieved relevant background knowledge to help explain the current sentence or idea. Other think-aloud processes coded, but not included in the focal analysis, included text rehearsal, predictive inferences, associations, monitoring, affective responses, non-responses, and an "other" category. See Table 4 for definitions and examples of participant responses for each think-aloud category. Interrater agreement was 89% (Fleiss' *kappa* = .75). Any disagreements between the raters were resolved via discussion.

The number of times participants engaged in each think-aloud process was tabulated. Because each participant engaged in a different number of processes, the overall proportion with which participants engaged in each process was computed (number of responses for a specific process divided by the total number of processes produced by the participant). For example, if a participant elaborated two times, but generated a total number of 23 processes, the proportion with which they elaborated would be 8.70% (i.e., 2/23). A number of processes occurred in less than 1% of the responses and were removed from analyses (a loss of 1.64% of the data, including predictive inferences (.10%), questions (.10%), associations (.70%), affective responses (.70%), and no response (.04%). Data from four participants were also removed from the analysis. One reported being on medication causing drowsiness, one was a non-native English speaker, one reported a learning disability, and one experienced loud construction noises during the session that interfered with reading.

We ran Pearson correlations between the processes (see Table 5). With respect to the focal processes, elaborations negatively correlated with backward inferences ($r = -.28$, $p < .05$). The correlations indicated no statistically significant positive relations between the remaining variables, suggesting each processing category was independent.

Coding of recall data. The recall transcripts were parsed and coded as in Experiment 1. The experimental conditions of the participants were masked to the raters. The percentage agreement among raters was 93% (Fleiss' *kappa* = .81).

Results

As in Experiment 1, we conducted separate 3 x 2 mixed model univariate analyses of variance (ANOVA) with belief/task-alignment (the combination of task instructions and pre-existing beliefs: match, mismatch, or neutral) as a between-subjects variable, and segment type (task-relevant or task-irrelevant) as a within-subject variable on the recall and think-aloud data. We again used the Greenhouse-Geisser adjustment. We computed partial eta squared (η^2) for the measurement of effect size. All follow-up post-hoc tests used Tukey's HSD, which controls for family wise error at the .05-level. We only report post-hoc tests that were significant at the $p <$ .05 level of significance unless otherwise noted.

Think-alouds. Table 6 contains descriptive statistics[2] related to the think-aloud data. These data are reported as the proportion with which participants engaged in each process.

[2] We ran analyses with the non-focal processes. For text rehearsal, the main effect of text segment was approaching significance, $F(1, 73) = 2.95$, $p = .09$, $\eta^2 = .04$. Participants engaged in numerically more text rehearsal for task-

Evaluations. The main effects for segment type [$F(1, 73) = .31$, $p = .58$, $\eta^2 = .004$] and belief/task-alignment [$F(2, 73) = .35$, $p = .70$, $\eta^2 = .01$] were not significant, nor was the interaction significant [$F(2, 73) = .55$, $p = .58$, $\eta^2 = .01$]. However, to gain a better understanding of readers' use of evaluations, we coded each evaluation as being supportive, refutational, or impartial toward the content of the segment. Then, we computed the proportional use of each type of evaluation for each participant (e.g., the number of supportive evaluations generated was divided by the total number of evaluations generated[3]). We ran a 2 (segment type: task-relevant or task-irrelevant) x 3 (belief/task-alignment: match, mismatch, neutral) for each type of evaluation. Examining refutational versus supportive evaluations enabled us to directly test for possible belief-consistency patterns. That is, given that particular text segments were either task-relevant or task-irrelevant, it was possible to investigate whether these segments were processed differently as a function of topic beliefs. For instance, do readers process task-relevant segments differently when they are belief-consistent versus when they are belief-inconsistent?

Supportive Evaluations. The main effects for text segment, $F(1, 73) = 8.94$, $p = .004$, $\eta^2 = .11$, and belief/task-alignment were significant, $F(1, 73) = 3.99$, $p < .05$, $\eta^2 = .10$. However, these main effects were qualified by the significant interaction between text segment and belief/task-alignment, $F(2, 73) = 22.69$, $p < .001$, $\eta^2 = .38$ (see Figure 2). Post-hoc tests indicated that readers used supportive evaluations for text segments differently based on belief-task alignment. When the text segment was task-relevant and there was a match between belief/task-alignment (i.e., task-relevant segments were belief-consistent), participants provided

relevant than task-irrelevant text. All other main effects and interactions for text rehearsal, monitoring, and the "other" category were not significant (F's ≤ 1.36, p's $\geq .28$).

[3] We divided by the total number of evaluations generated because we were interested in examining different subtypes of evaluations. However, we also ran the analyses with the denominator as the total number of think aloud processes in general, and obtained similar results.

more supportive evaluations for task-relevant segments (belief-consistent) than for task-irrelevant segments (belief-inconsistent; $p < .01$). However, when the text segment was task-relevant and there was a mismatch between belief/task-alignment (i.e., task-relevant segments were belief-*inconsistent*), participants provided fewer supportive evaluations for task-relevant segments (belief-inconsistent) than for task-irrelevant segments (belief-consistent; $p < .01$). Further, when a text segment was task-relevant, and the reader had neutral beliefs (belief/task-neutral), participants provided more supportive evaluations for task-relevant segments than for task-irrelevant segments ($p < .01$).

For task-relevant text, the belief/task-match group provided more supportive evaluations than the other groups (p's $< .01$), and the belief/task-neutral group provided more supportive evaluations than the belief/task-mismatch group ($p < .01$). For task-irrelevant text, the belief/task-mismatch group provided more supportive evaluations than the other groups (p's $<$ .01).

Refutational Evaluations. Neither the main effect for text segment $F(1, 73) = 3.59$, p = .06, $\eta^2 = .05$, nor the main effect for belief/task-alignment, $F(2, 73) = .27$, $p = .76$, $\eta^2 = .01$, were significant. However, the interaction was significant, $F(2, 73) = 19.73$, $p < .001$, $\eta^2 = .35$ (as illustrated in Figure 3). Post-hoc tests indicated that readers used refutational evaluations for text segments differently based on belief-task alignment. When the text segment was task-relevant and there was a match between belief/task-alignment (i.e., task-relevant segments were belief-consistent), participants provided fewer refutational evaluations for task-relevant segments (belief-consistent) than for task-irrelevant segments (belief-inconsistent; $p < .01$). However, when the text segment was task-relevant, and there was a mismatch between belief/task-alignment (i.e., task-relevant segments were belief-*inconsistent*), participants provided more refutational evaluations for task-relevant segments (belief-inconsistent) than for task-irrelevant segments (belief-consistent; $p < .01$). Further, when a text segment was task-relevant, and the reader had neutral beliefs (belief/task-neutral), participants provided fewer refutational evaluations for task-relevant segments than for task-irrelevant segments ($p < .01$).

For task-relevant text, the belief/task-mismatch group (for whom the task-relevant text was belief-inconsistent) provided more refutational evaluations than the other groups (p's < .01). For task-irrelevant text, the belief/task-mismatch group (for whom the task-relevant text was belief-consistent) provided fewer refutational evaluations than the other groups (p's < .01).

Impartial Evaluations. The main effect of text segment was significant, $F(1, 73) = 4.24$, $p < .05$, $\eta^2 = .05$. Participants used more impartial evaluations for task-irrelevant than task-relevant segments. The main effect of belief/task-alignment was significant, $F(1, 73) = 3.24$, $p < .05$, $\eta^2 = .08$. The belief/task-neutral group provided more impartial evaluations than the belief/task-match group ($p = .05$) and numerically more than the belief/task-mismatch group ($p = .11$). The interaction was not significant, $F(2, 73) = .44$, $p = .65$, $\eta^2 = .01$.

Backward Inferences. Neither the main effects for text segment [$F(1, 73) = .28$, $p = .60$, $\eta^2 = .004$] nor belief/task-alignment were significant [$F(2, 73) = .27$, $p = .77$, $\eta^2 = .01$]. However, the interaction between text segment and belief/task-alignment was significant (as illustrated in Figure 4), $F(2, 73) = 3.53$, $p < .05$, $\eta^2 = .09$. When the text was task-irrelevant, the belief/task-match group (for whom the task-irrelevant text was belief-inconsistent) provided more backward inferences than the belief/task-mismatch group (for whom the task-irrelevant text was belief-consistent) or for the belief/task-neutral group (p's < .01).

Elaborations. The main effect of text segment was significant, $F(1, 73) = 5.33$, $p < .05$, $\eta^2 = .07$. Participants provided more elaborations for task-irrelevant text segments than for task-relevant text segments. The main effect of belief/task-alignment was not significant, $F(2, 73) = .80$, $p = .45$, $\eta^2 = .02$, nor was the interaction, $F(2, 73) = 1.59$, $p = .21$, $\eta^2 = .04$.

Recall. Table 7 contains descriptive statistics related to the recall data, which are reported as the proportion of unique text ideas recalled. The main effect of text segment was significant, $F(1, 73) = 7.98$, $p < .01$ $\eta^2 = .10$. Participants recalled task-relevant segments to a greater extent than task-irrelevant segments. The main effect of belief/task alignment was significant, $F(2, 73) = 3.04$, $p = .05$, $\eta^2 = .08$. Participants in the belief/task-neutral group recalled more segments than participants in the belief/task-mismatch group ($p < .05$). No other post-hoc comparisons were significant. The interaction was not significant, $F(2, 73) = 1.37$, $p = .26$, $\eta^2 = .04$.

Discussion

For the think-alouds, readers' beliefs influenced online processes more than task instructions. Students in the belief/task-match group used supportive evaluations more than refutational evaluations for task-relevant text (which was belief-consistent), whereas they used refutational evaluations more than supportive evaluations for task-irrelevant text (which was belief-inconsistent). Students in the belief/task-mismatch group used refutational evaluations more than supportive evaluation for task-relevant text (which was belief-inconsistent), whereas they used supportive evaluations more than refutational evaluations for task-irrelevant text (which was belief-consistent). Thus, readers' use of evaluations was driven by whether the text was belief-consistent rather than whether the text was task-relevant.

Further, students in the belief/task-match group were more likely to generate backward inferences for task-irrelevant text (which was belief-inconsistent) text than the other readers, lending partial support for a disconfirmation strategy. These data suggest that individuals sought to both confirm belief-consistent text and disconfirm belief-inconsistent text. Nonetheless, participants were more likely to generate elaborations for task-irrelevant text than for task-relevant text independently of whether the information was belief-consistent.

For the recall data, task-relevant information was recalled better than task-irrelevant information across all groups. This is consistent with the results from Experiment 1, and suggests that task instructions affected memory for text, independently of beliefs. However, beliefs still played a role in recall. Participants in the belief/task-neutral groups recalled more text than participants asked to focus on content inconsistent with their beliefs (the belief/task-mismatch group).

Conclusions and General Discussion

We used the three-pronged approach (Magliano & Graesser, 1991) to investigate the effects task instructions and topic beliefs on moment-by-moment processing of and memory for belief-related text. This approach allowed us to examine the association between online reading processes and memory (Rapp & van den Broek, 2005), as well as how reading time data can potentially mask differential processing. In Experiments 1 and 2, participants recalled task-relevant information better than task-irrelevant information independently of whether the information was belief-consistent. However, Experiment 1 showed that there were no differences in reading times as a function of belief/task-alignment, whereas Experiment 2 showed differences in the nature of reading processes as a function of belief/task-alignment. When participants held stronger beliefs, topic beliefs overrode task instructions during reading, whereas participants who held neutral beliefs enacted task instructions in a more uniform manner. The findings suggest that topic beliefs exerted a greater influence on online processing of belief-related text, whereas task instructions exerted a greater influence on memory for belief-related text.

The reading time, think-aloud, and recall data each provided different information. Although reading times were not affected by beliefs or task instructions, think-aloud processing varied as a function of beliefs, such that participants appeared motivated to protect their beliefs. For instance, refutational evaluations were used more than supportive evaluations at belief-inconsistent text, whereas supportive evaluations were used more than refutational evaluations at belief-consistent text, independently task instructions. Thus, although reading times for belief-consistent and belief-inconsistent segments did not differ, the think-aloud data indicated differences in the processes readers used while reading these segments.

When readers had neutral beliefs, they provided more supportive evaluations than refutational evaluations when they read task-relevant text, and more refutational evaluations than supportive evaluations when they read task-irrelevant text. This suggests that task instructions may have affected the valence of the information (McCrudden et al., 2010). In addition, all participants provided more elaborations for task-irrelevant text segments than for task-relevant text segments. If the task instructions indeed affected the valence of the information, the elaboration of task-irrelevant text segments is consistent with previous research in which individuals more carefully scrutinize belief-inconsistent arguments (i.e., disconfirmation bias; Edwards & Smith, 1996).

However, recall varied as a function of task instructions, such that participants recalled task-relevant information to a greater extent than task-irrelevant information, independently of whether the information was belief-consistent. This suggests that the expectation that task-relevant information would be needed on the post-reading task facilitated recall of this information. This finding is consistent with research which has shown that reader expectations about post-reading assessment can affect memory, even when asked to focus on the same information (McNamara & Dempsey, 2011; van den Broek et al., 2011).

Although previous research has shown that readers generally spend more time reading task-relevant text and recall more of this information than task-irrelevant text, this is not always the case (McCrudden, Schraw, & Kambe, 2005; Rapp & Mensink, 2011; Rapp & van den Broek, 2005). Although task instructions may affect moment-by-moment processing, they do not necessarily affect memory for text. For instance, while readers might focus on task-relevant information during reading, this does not necessarily prevent readers from processing or encoding task-irrelevant information. Similarly, although task instructions may affect memory for text, they do not necessarily affect moment-by-moment processing. For instance, the value of information that is encountered during reading may change after reading when it is considered in relation to the entire text.

Task instructions can differ in their specificity, ranging from more specific to more general (Goldman & Durán, 1988; McCrudden & Schraw, 2007). Maier and Richter (2015) asked readers to approach a text with the general goals of summarizing or building an argument. Such general instructions may be more open to interpretation, allowing beliefs to play a greater role in online processes and memory. In contrast, the present study utilized specific relevance instructions that are more explicit and less open for interpretation. The specific relevance instructions led participants to include more task-relevant textual ideas in their recalls, independently of whether the ideas were belief-consistent. Thus, the present study adds to the growing body of research by indicating that when asking students to comprehend material they do not agree with, teachers may need to provide specific instructions to focus on belief-inconsistent text (Maier & Richter, 2014). In addition, the present study provides some evidence that inconsistencies in prior work may be explained by how moderating variables (such as task instructions) can modify when and how beliefs affect memory. Hence, future work should clearly identify potential moderating variables for understanding when and how beliefs can affect processing and memory.

Another potential moderating variable might be the type of text with which readers interact. The present study utilized extended, connected discourse, which differs from studies utilizing thought-listing tasks or single-sentence arguments. Extended, connected text is more like the types of texts students encounter in educational settings, and combining that with a think-aloud methodology provided us the ability to assess overt cognitive strategies in relation to offline recall.

Limitations and Future Directions for Research

There are several directions for future research. Reading times are useful for investigating what readers pay attention to during reading. However, the self-paced, phrase-by-phrase mode of reading does not mimic naturalistic reading, in which individuals can re-read previous information and view entire sentences or paragraphs. Think-alouds provide detailed information about the processes readers engage in during reading. However, verbalizing one's thoughts during reading may change how readers process text (Magliano & Graesser, 1991; Nisbett & Wilson, 1977). Therefore, future research could investigate online processes that more closely reflect naturalistic reading, such as with the use of eye-tracking, which can provide a precise measure of attention allocation. Similarly, it is possible to measure online process and offline products in a first experiment, followed by a second experiment in which participants read silently from paper and then complete offline measures. This approach would be useful for determining the extent to which the offline measures show similar patterns when reading under more- and less-naturalistic settings.

Particular types of post-reading tasks may moderate the effects of beliefs on memory. For example, Maier and Richter (2013) found that participants recalled belief-inconsistent text at the textbase level of representation, whereas belief-consistent text was better integrated into the reader's situational model level of representation (which contains inferences and connections to prior knowledge; Kintsch & van Dijk, 1978). In the present study, the recall task did not distinguish between the textbase versus situation model levels of representation. Although it is notable that the relevance instructions fostered memory independently of beliefs, future work should examine the extent to which belief- versus task-relevant information is encoded along varying levels of representation, and for different types of text.

In the present study, we measured readers' topic beliefs. Future research could investigate the effects of additional individual difference variables on online processes and offline products. For instance, additional individual difference variables that could be investigated include prior knowledge, reading abilities, working memory, emotion, or need for cognition. Another avenue of future research may be to examine the extent to which the direction of one's beliefs potentially mediates processing.

The arguments in the text were plausible and attributed to credible sources. Nonetheless, we did not collect participant ratings on the perceived quality of the arguments or the credibility of the sources. Thus, an interesting avenue for future research would be to investigate the extent to which argument quality and source credibility affect processing and memory.

Finally, the study had some limitations that merit consideration. Regarding the belief ratings, we used a one-item scale that directly addressed the specific topic of interest (i.e., teaching ID in science classrooms). However, it would be useful for future work to include a belief measure with multiple items to increase reliability. In addition, the goal of the study was to examine potential interactions between the task, beliefs, and text content. As such, we designed the study to include categorical assignment to task instructions, categorical text segments (relevant vs. irrelevant to the task instructions), and crossing belief categories with task instructions. This design facilitated the examination of interactions. However, we did not have the statistical power to examine beliefs and task instructions as separate variables. Future work might consider larger sample sizes and utilizing continuous variables to address potential non-linear relations and account for strengths of beliefs. As one final consideration, the recall task allowed us to quantify how many unique relevant versus irrelevant ideas participants included. Although we attempted to minimize demand characteristics by providing general instructions and not reminding participants of their reading goal, it is still possible participants purposefully omitted information that did not align with the task instructions, and hence the inclusion of smaller amounts of text. Future work should compare different types of memory tasks.

In practical classroom settings, students must learn about belief-related topics, yet beliefs can affect reading processes and memory for text (Maier & Richter, 2014). Thus, teachers should consider providing appropriate instructional supports and strategies when students are expected to build knowledge about belief-related topics (Sinatra et al., 2014). Task instructions can be used to help students recall belief-related information. The ability to recall information may be a first step in helping students understand different viewpoints and developing a strong knowledge base in a content area. Future research should investigate ways to further improve understanding of conflicting viewpoints. For example, writing tasks could encourage students to demonstrate understanding by summarizing key ideas and arguments for different sides of an issue, and then have students argue in favor of and in opposition to these viewpoints. This may help students make a clear distinction between what they are expected to understand and what they believe.

References

Alexander, P. A., & Jetton, T. L. (2000). Learning from text: A multidimensional and developmental perspective. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research, Vol. III* (pp. 285-310). Mahwah, NJ: Lawrence Earlbaum.

Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology, 39*(3), 206-219. doi: 10.1016/j.cedpsych.2014.05.003

Coté, N., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In H. Van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 169-193). Mahwah, NJ: Lawrence Erlbaum Associates.

E-Prime. (Version 2.0) [Computer software]. Sharpsburg, PA: Psychological Software Tools, Inc.

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

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

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: The MIT Press.

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Summary versus argument tasks when working with multiple documents: Which is better for whom? *Contemporary Educational Psychology, 35*, 157-173.

Goldman, S. R., & Durán, R. P. (1988). Answering questions from ocenography texts: Learner, task, and text characteristics. *Discourse Processes, 11*(4), 373-412. doi: 10.1080/01638538809544710

Goldman, S. R., & Varma, S. (1995). CAPping the construction-integration model of discourse comprehension. In I. Charles A. Weaver, S. Mannes & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 337-358). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Hart, W., Eagly, A. H., Lindberg, M. J., Albarraccin, D., Brechan, I., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin, 135*(4), 555-588.

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

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). Charlotte, NC: Information Age Publishing.

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2002). Perspective effects on online text processing. *Discourse Processes, 33*(2), 159-173. doi: 10.1207/s15326950dp4003_4

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2003). How prior knowledge, WMC, and relevance of information affect eye fixations in expository text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(3), 447-457.

Kendeou, P., Bohn-Gettler, C. M., & Fulton, S. (2011). What we have been missing: The role of goals in reading comprehension. In M. T. McCrudden, J. P. Magliano & G. Schraw (Eds.), *Text Relevance and Learning from Text* (pp. 375-394). Greenwich, CT: Information Age Publishing.

Kendeou, P., Bohn-Gettler, C. M., White, M. J., & van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading, 31*(3), 259-272. doi: 10.1111/j.1467-9817.2008.00370.x

Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition, 35*(7), 1567-1577.

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

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394. doi: 10.1037/0033-295X.85.5.363

Klaczynski, P. A., & Robinson, B. (2000). Personal theories, intellectual ability, and epistemological beliefs: Adult age differences in everyday reasoning biases. *Psychology and Aging, 15*(3), 400-416. doi: 10.1037//0882-7TO.15.3.400

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged approach for studying inference generation in literary text. *Poetics, 20*, 193-232.

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction, 21*(3), 251-283.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91*(4), 615-629.

Maier, J., & Richter, T. (2013). Text belief consistency effects in the comprehension of multiple texts with conflicting information. *Cognition and Instruction, 31*(2), 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*, 51-74. doi: 10.1007/s11409-013-9111-x

Maier, J., & Richter, T. (2015). Effects of text-belief consistency and reading task on the strategic validation of multiple texts. *European Journal of Psychology of Education, Online First*, 1-19. doi: 10.1007/s10212-015-0270-9

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals, and text processing: A mixed methods study. *Contemporary Educational Psychology, 35*(4), 229-241. doi: 10.1016/j.cedpsych.2009.12.001

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2011). *Text relevance and learning from text*. Charlotte, NC: Information Age Publishing.

McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review, 19*(2), 113-139. doi: 10.1007/s10648-006-9010-7

McCrudden, M. T., & Schraw, G. (2010). The effects of relevance instructions and verbal ability on text processing. *Journal of Experimental Education, 78*(1), 96-117.

McNamara, D. S., & Dempsey, K. (2011). Reader expectations of question formats and difficulty: Targeting the zone. In M. T. McCrudden, J. P. Magliano & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 321-352). Charlotte, NC: Information Age Publishing.

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

Moshman, D. (2009). *Liberty and Learning*. Portsmouth, NH: Heinemann.

Murphy, P. K., & Mason, L. (2006). Changing knowledge and beliefs. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 305-324). Mahwah, NJ: Lawrence Erlbaum.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 8*, 231-259. doi: 10.1037/0033-295X.84.3.231

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology*, *25*(3), 241-286. doi: 10.1006/ceps.2000.1040

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Erlbaum.

Rapp, D. N., & van den Broek, P. (2005). Dynamic text comprehension: An integrative view of reading. *Current Directions in Psychological Science*, *14*(5), 276-279. doi: 10.1111/j.0963-7214.2005.00380.x

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

Schraw, G., Wade, S. E., & Kardash, C. M. (1993). Interactive effects of text-based and task-based importance on learning from text. *Journal of Educational Psychology, 85*, 652-661.

Sinatra, G. M., Kienhues, D., & Hofer, B. K. (2014). Addressing challenges to public understanding of science: Epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist, 49*, 123-138.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Poligical Science*, *50*(3), 755-769. doi: 10.1111/j.1540-5907.2006.00214.x

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes, 21*, 255-288.

van den Broek, P., Bohn-Gettler, C. M., Kendeou, P., Carlson, S., & White, M. J. (2011). The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. P. Magliano & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 123-139). Charlotte, NC: Information Age Publishing.

van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition, 29*, 1081-1087. doi: 10.3758/BF03206376

van den Broek, P., Risden, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch & E. J.

O'Brien (Eds.), *Sources of Coherence in Reading* (pp. 353-373). Hillsdale, NJ: Lawrence Erlbaum.

Vidal-Abarca, E., Mañá, A., & Gil, L. (2010). Individual differences for self-regulating task-oriented reading. *Journal of Educational Psychology*, *102*, 817-826.

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

Wolfe, M. B. W., Tanner, S. M., & Taylor, A. R. (2013). Processing and representation of arguments in one-sided texts about disputed topics. *Discourse Processes*, *50*(7), 457-497. doi: 10.1080/0163853X.2013.828480

Table 1

Sample sizes of participants receiving focus instructions as aligned with pre-existing beliefs

Participants' Pre-Existing Beliefs	Experiment 1: Task Instructions		Experiment 2: Task Instructions	
	For	Against	For	Against
For	12	17	19	16
Against	16	13	7	12
Neutral	17	13	12	10

Table 2

Experiment 1: Reading Times Per Word (in milliseconds) by Group

		Text Segment							
		Relevant		Irrelevant		Overall		Relevant - Irrelevant	
Belief/Task Alignment	*n*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*
Match	25	348.6 (16.1)	[316.6, 380.5]	349.6 (18.0)	[313.8, 385.4]	349.0 (16.6)	[316.2, 381.8]	-1.0 (10.1)	[-21.9, 19.8]
Mismatch	33	317.6 (14.0)	[289.8, 345.4]	305.5 (15.7)	[274.3, 336.7]	312.7 (14.4)	[284.1, 341.2]	12.1 (8.2)	[-4.6, 28.9]
Neutral	30	353.2 (14.7)	[324.0, 382.4]	347.3 (16.4)	[314.6, 380.0]	349.4 (15.1)	[319.5, 379.4]	5.9 (8.9)	[-12.4, 24.2]
Overall	88	339.8 (8.6)	[322.7, 356.9]	334.1 (9.7)	[314.9, 353.3]	337.0 (8.7)	[319.6, 354.4]	5.7 (5.1)	[-3.8, 16.5]

Table 3

Experiment 1: Overall Proportion of Text Recalled by Each Group

		Text Segment							
		Relevant		Irrelevant		Overall		Relevant - Irrelevant	
Belief/Task Alignment	*n*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*
Match	25	.09 (.01)	[.07, .11]	.04 (.01)	[.02, .06]	.09 (.01)	[.06, .08]	.05 (.01)	[.02, .08]
Mismatch	33	.07 (.01)	[.05, .09]	.03 (.01)	[.02, .05]	.07 (.01)	[.06, .08]	.04 (.01)	[.02, .05]
Neutral	30	.07 (.01)	[.05, .09]	.05 (.01)	[.03, .06]	.07 (.01)	[.06, .09]	.02 (.01)	[.003, .05]
Overall	88	.08 (.005)	[.07, .09]	.04 (.005)	[.03, .05]	.06 (.004)	[.05, .07]	.04 (.01)	[.02, .05]

Table 4

Experiment 2: Definitions and Examples of Each Think-Aloud Process

Process	Definition	Text Excerpt	Sample Participant Response
Evaluations	Stating opinions about the content of the text, further coded as supportive, refutational, or impartial		
	Supportive Evaluations	Those who are opposed to teaching ID in science classes argue that if alternatives to scientific explanations for the diversity of life are taught, it should *not* take place in science classes.	"Um, again I agree that it shouldn't take place in a science class."
	Refutational Evaluations	Those who are in favor of teaching ID in science classes argue that it is counter-productive to exclude non-scientific alternatives from science lessons because it alienates some children from science.	"Non-scientific alternatives do not really belong in a scientific…class."
	Impartial Evaluations	Either way, it seems that the topic of whether ID should be taught in science classes will continue to be a controversial issue.	"I think both sides have, um, good viewpoints."
Backward Inferences	Referring to information presented in earlier sentences that are relevant to, or explain, the current sentence	Those who are in favor of teaching ID argue that teaching ID in science classes can help students overcome their misconceptions about science and learn to think critically about science.	"You can't take data, um, you can't replicate the experiment." (Referencing 4 sentences earlier: *In science, explanations are restricted to results obtained through observations and experiments that can be substantiated by other researchers or scientists.*)
Elaborations	Retrieving relevant background knowledge to help explain the current sentence or idea	Those who are opposed to teaching ID in science classes argue that if alternatives to scientific explanations for the diversity of life are taught, it should *not* take place in science classes.	"Obviously like with budget cuts and everything, they don't really have, um, the money to teach it in a separate class. of earth, um, came about."
Text Rehearsal	Paraphrases or repetitions of the text that captured the gist meaning of the sentence	Those who are opposed to teaching ID in science classes argue that if alternatives to scientific explanations for the diversity of life are taught, it should *not* take place in science classes.	**"**Um… Science—alternatives to scientific explanations uh, should not take place in scientific c---science classes."
Predictive Inferences	Anticipations about what will come next, or logical outcomes based on the text	Either way, it seems that the topic of whether ID should be taught in science classes will continue to be a controversial issue.	"And that's going to continue till one side proves they're right either way."
Associations	Retrieving background knowledge that is not relevant to the topic	Those who are opposed to teaching ID in science classes argue that ID is simply not science.	"Darwin] Guy [Darwin] was a nut case… and an anti-Semite."
Monitoring	Statements reflecting on one's own understanding	Those who are in favor of teaching ID argue that teaching ID in science classes can help students overcome their misconceptions about science and learn to think critically about science.	"Um, I don't really know how that would work."
Questions	Asking a question of the experimenter or about the text	Those who are opposed to teaching ID in science classes argue that ID does not stimulate students or researchers to pursue knowledge.	"Do you like going to class?"
Affective Responses	Making a comment about one's own emotions, or having an emotional reaction (i.e., laughing)	Next, you will read some arguments from some people who are in favor and from some who are opposed to teaching ID in science classrooms.	[laughs]
Non-Responses	Verbally stating that he/she does not have a response, or not responding to a starred sentence	Those who are in favor of teaching ID in science classes argue that it is counter-productive to exclude non-scientific alternatives from science lessons because it alienates some children from science.	"I have no comment on that."
Other	Any other response that did not fall into any of the other categories	Those who are opposed to teaching ID in science classes argue that ID does not stimulate students or researchers to pursue knowledge.	"I like cats."

Table 5

Experiment 2: Descriptive Statistics and Pearson Correlations for the Proportion of Think-Aloud Processes (n = 76)

Process	Mean	SD	1	2	3	4	5	6
1. Text Rehearsal	.10	.16	-	.11	-.44***	-.47***	-.16	-.10
2. Backward Inferences	.06	.08		-	-.28*	-.20	-.14	-.19
3. Elaborations	.38	.18			-	-.22	-.28*	-.28**
4. Evaluations	.35	.14					-.13	-.03
5. Monitoring	.07	.09					-	.27*
6. Other	.03	.07						-

*$p < .05$; **$p < .01$; ***$p < .001$

Table 6

Adjusted Means (and Standard Errors) for Proportions of Think Aloud Processes

			Text Segment							
			Relevant		Irrelevant		Overall		Relevant - Irrelevant	
Process	Belief / Task Alignment	*n*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*
Evaluations	Match	31	.35 (.04)	[.28, .43]	.34 (.03)	[.28, .40]	.35 (.03)	[.31, .41]	.01 (.04)	-.07, .09]
	Mismatch	23	.31 (.04)	[.22, .39]	.32 (.03)	[.26, .39]	.33 (.03)	[.26, .39]	-.02 (.04)	[-.10, .06]
	Neutral	22	.37 (.04)	[.28, .45]	.32 (.03)	[.25, .39]	.35 (.03)	[.28, .42]	.05 (.05)	[-.06, .16]
	Overall	76	.34 (.02)	[.30, .39]	.33 (.02)	[.29, .37]	.34 (.02)	[.30, .37]	.01 (.02)	[-.04, .06]
Supportive Evaluations (percentage within all evaluations)	Match	31	.80 (.06)	[.69, .91]	.28 (.06)	[.17, .40]	.43 (.03)	[.36, .49]	.52 (.08)	[.36, .67]
	Mismatch	23	.26 (.07)	[.13, .39]	.63 (.07)	[.50, .76]	.38 (.04)	[.30, .46]	-.37 (.12)	[-.62, -.12]
	Neutral	22	.57 (.07)	[.44, .71]	.20 (.07)	[.06, .34]	.33 (.04)	[.26, .41]	.37 (.11)	[.14, .60]
	Overall	76	.54 (.04)	[.47, .62]	.37 (.04)	[.30, .45]	.46 (.02)	[.41, .50]	.17 (.07)	[.06, .35]
Refutational Evaluations (percentage within all evaluations)	Match	31	.16 (.06)	[.05, .28]	.63 (.06)	[.51, .76]	.33 (.03)	[.27, .39]	-.47 (.08)	-.64, -.30]
	Mismatch	23	.63 (.07)	[.49, .76]	.21 (.07)	[.07, .36]	.33 (.03)	[.26, .40]	.42 (.12)	[.16, .67]
	Neutral	22	.23 (.07)	[.09, .36]	.52 (.07)	[.38, .67]	.29 (.04)	[.22, .36]	-.30 (.12)	[-.54, -.05]
	Overall	76	.34 (.04)	[.27, .41]	.46 (.04)	[.38, .54]	.40 (.02)	[.35, .45]	-.12 (.07)	[-.30, -.003]
Impartial Evaluations (percentage within all evaluations)	Match	31	.03 (.02)	[-.01, .08]	.05 (.03)	[-.01, .10]	.04 (.02)	[.002, .07]	-.01 (.02)	[-.05, .02]
	Mismatch	23	.02 (.02)	[-.03, .08]	.07 (.03)	[.01, .13]	.04 (.02)	[.003, .09]	-.05 (.02)	[-.10, .00]
	Neutral	22	.11 (.03)	[.06, .16]	.14 (.03)	[.07, .20]	.10 (.02)	[.06, .14]	-.04 (.03)	[-.10, .04]
	Overall	76	.06 (.01)	[.03, .08]	.09 (.02)	[.05, .12]	.07 (.01)	[.04, .10]	-.03 (.01)	[-.06, .00]
Backward Inferences	Match	31	.045 (.02)	[.01, .08]	.08 (.017)	[.05, .11]	.07 (.01)	[.04, .10]	-.03 (.02)	[-.07, .003]
	Mismatch	23	.061 (.02)	[.02, .10]	.032 (.02)	[-.01, .07]	.05 (.02)	[.01, .08]	.03 (.02)	[-.01, .07]
	Neutral	22	.067 (.02)	[.03, .11]	.044 (.02)	[.003, .08]	.05 (.02)	[.02, .09]	.02 (.02)	[-.02, .07]
	Overall	76	.06 (.01)	[.04, .08]	.05 (.01)	[.03, .07]	.05 (.01)	[.04, .07]	.01 (.01)	[-.02, .02]
Elaborations	Match	31	.40 (.04)	[.32, .49]	.40 (.03)	[.34, .47]	.40 (.03)	[.33, .46]	.001 (.02)	[-.05, .05]
	Mismatch	23	.35 (.05)	[.26, .45]	.42 (.04)	[.34, .49]	.39 (.04)	[.31, .47]	-.06 (.04)	[-.15, .02]
	Neutral	22	.30 (.05)	[.20, .40]	.37 (.04)	[.30, .45]	.35 (.04)	[.28, .43]	-.07 (.04)	[-.15, .00]
	Overall	76	.35 (.03)	[.30, .41]	.40 (.02)	[.35, .44]	.38 (.02)	[.33, .42]	-.04 (.02)	[-.08, -.001]

Note: Some totals in the Relevant – Irrelevant column are subject to rounding errors.

Table 7

Experiment 2: Overall Proportion of Text Recalled by Each Group

		Text Segment							
		Relevant		Irrelevant		Overall		Relevant - Irrelevant	
Belief/Task Alignment	*n*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*	*M (SE)*	95% *CI*
Match	31	.05 (.01)	[.04, .07]	.05 (.01)	[.03, .07]	.08 (.01)	[.07, .09]	.01 (.01)	[-.01, .02]
Mismatch	23	.05 (.01)	[.03, .08]	.03 (.01)	[.01, .05]	.07 (.01)	[.05, .08]	.03 (.01)	[.005, .05]
Neutral	22	.09 (.01)	[.06, .11]	.06 (.01)	[.03, .08]	.09 (.01)	[.08, .11]	.03 (.02)	[-.01, .07]
Overall	76	.07 (.01)	[.05, .08]	.04 (.01)	[.03, .06]	.05 (.005)	[.05, .06]	.02 (.07)	[.004, .03]

Note: Some totals in the Relevant – Irrelevant column are subject to rounding errors.

Figure 1

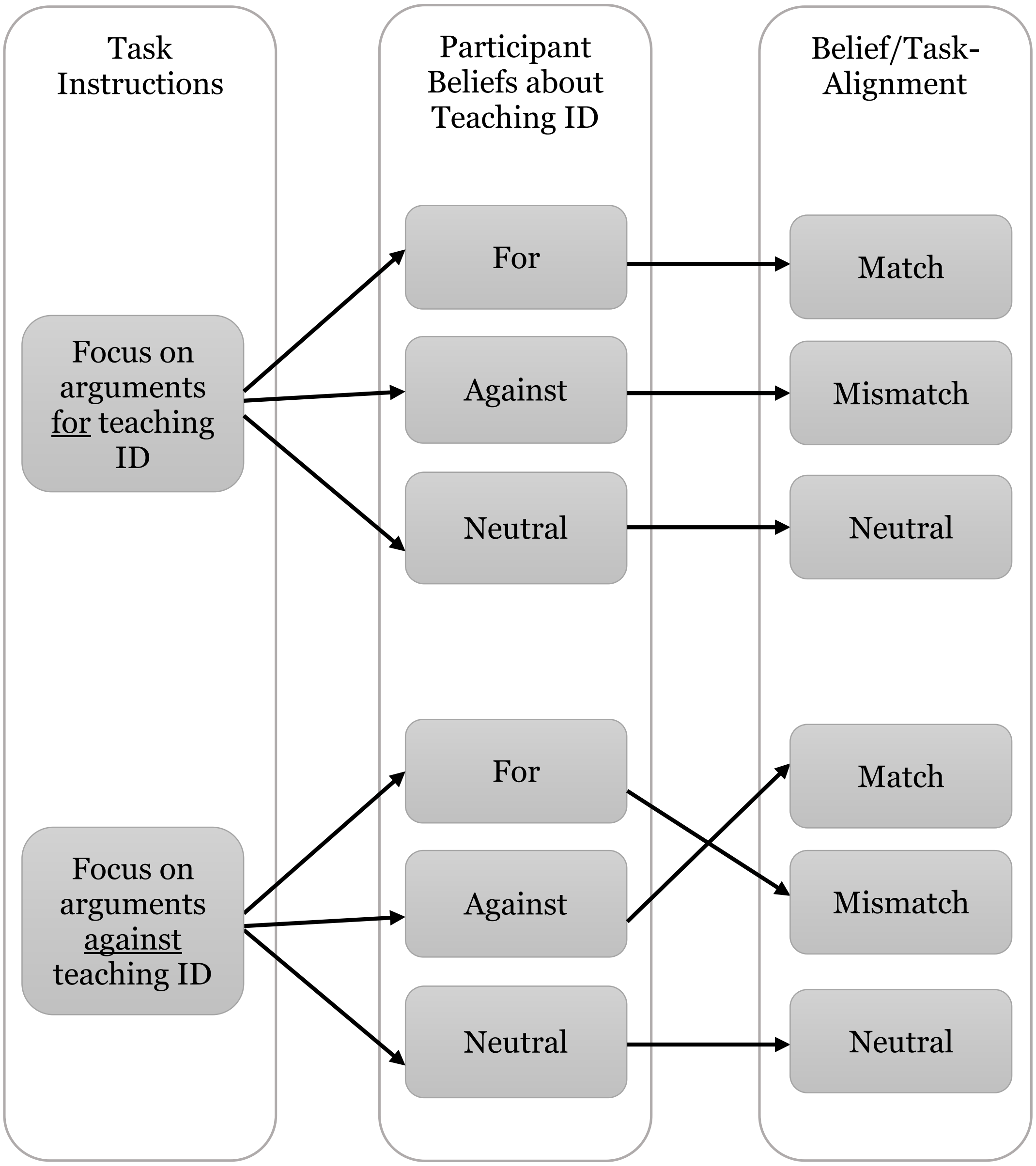
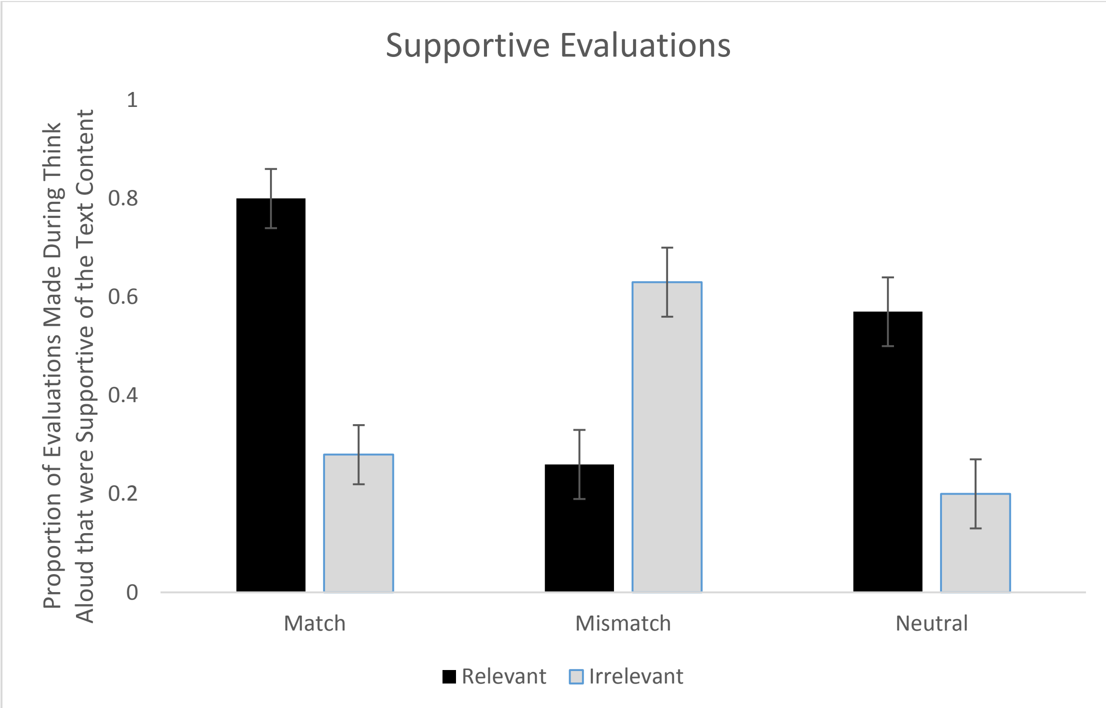
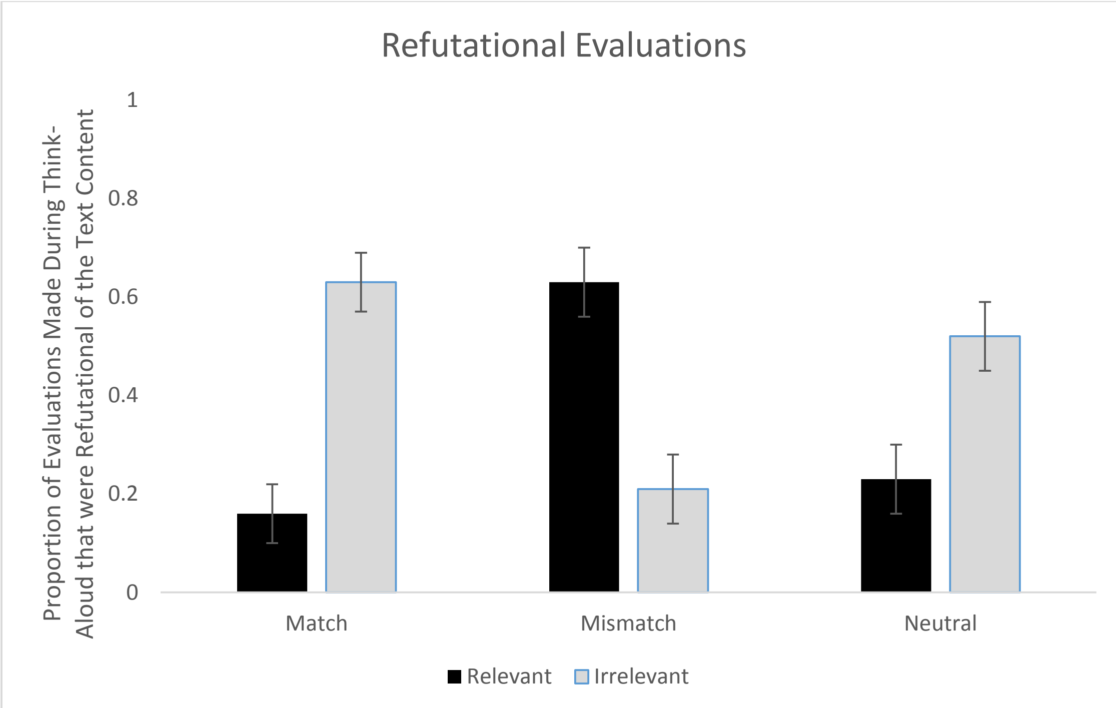

Task Instructions
Participant Beliefs about Teaching ID
Belief/Task-Alignment
Focus on arguments for teaching ID
For
Against
Neutral
Match
Mismatch
Neutral
Focus on arguments against teaching ID
For
Against
Neutral
Match
Mismatch
Neutral

Figure 2



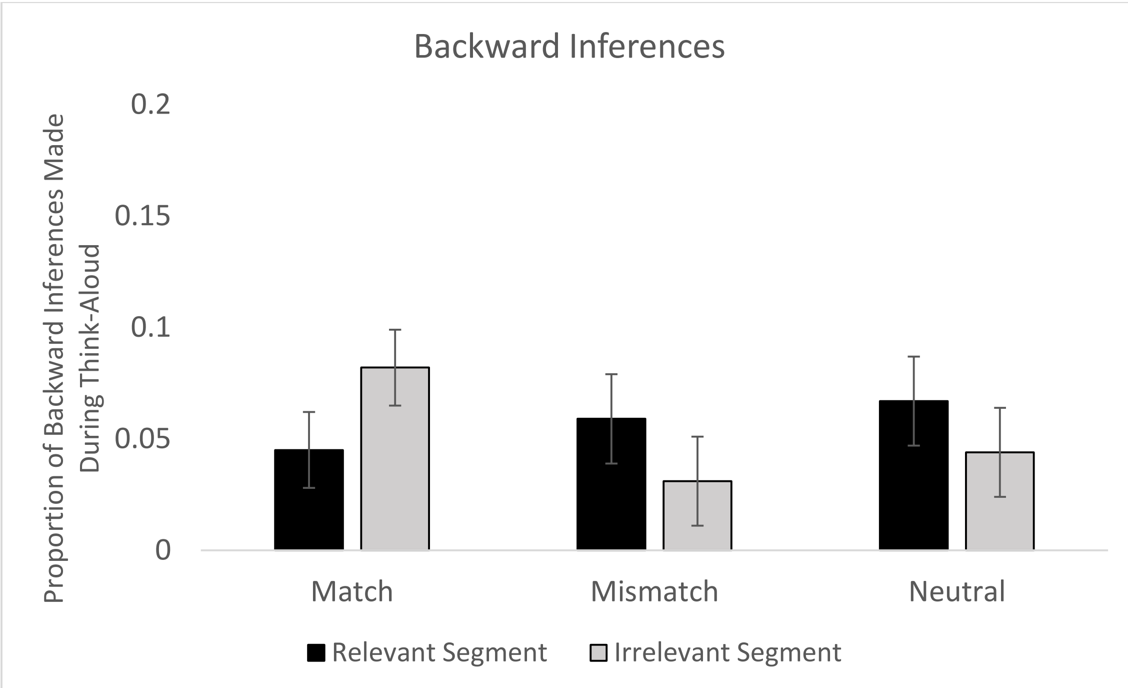
Note. Error bars represent standard errors.

Figure 3



Note. Error bars represent standard errors.

Figure 4



Note. Error bars represent standard errors.