

# Multiple-Text Comprehension and Evaluation: The Influence of Reading Goal, Belief Consistency, and Argument Type

**Sylvia M. Savvidou**

**Irene-Anna Diakidou**

*Department of Psychology, University of Cyprus, Nicosia, Cyprus*

**Lucia Mason**

*Department of Developmental Psychology and Socialization, University of Padova, Padova, Italy*

## ABSTRACT

The present study examined how argument type (science based vs. personal case based), belief consistency (belief consistent vs. inconsistent) and reading goals (read to evaluate vs. read to learn) influence comprehension and trustworthiness evaluations for claim-conflicting multiple texts. Undergraduates read four conflicting texts about the effects of vegan nutrition and completed four corresponding single-text comprehension and trustworthiness tasks before completing a multiple-text comprehension task. The results indicated better memory for personal case-based texts that capitalized on everyday life experiences and emotions than science-based texts in the multiple-text comprehension task. Reading to evaluate benefitted memory only for the belief-inconsistent personal text and contributed to lower trustworthiness ratings for all texts in comparison to reading to learn. The present study's findings highlight the importance of factors pertaining to argument quality, namely argument type, in comprehension and trustworthiness judgments.

Nowadays, with the aid of the Internet, people encounter multiple argumentative texts that are often claim-conflicting (Abendroth & Richter, 2020). People read multiple texts in a variety of contexts and for a variety of reasons: to learn, to form opinions, and to make important decisions (Betsch et al., 2011). When reading, multiple factors come into play and interact in a complex way. Based on the RAND model (RAND Reading Study Group & Snow, 2002), reading is influenced by readers' characteristics, the text itself, the reading activity, and the context. Reader characteristics can be cognitive (e.g., prior knowledge, working memory) or motivational (e.g., self-efficacy) and possibly interact with text-specific features, such as text genres (e.g., expository vs. narrative vs. persuasive) or different media forms (Coiro, 2021; RAND Reading Study Group & Snow, 2002). Reading activity includes reading goals (reading purposes, reading tasks) that can be externally or internally generated and, also, reading processing and outcomes, such as comprehension (of single or multiple texts) and learning. Especially, in the multiple-text context, an important reading activity is also the evaluation of sources and texts due to the vast amount of conflicting information and quality level in a set of multiple documents (Coiro, 2021). In the current study, we focused on belief consistency as reader characteristic, argument type as text feature, reading goals, trustworthiness evaluations, and comprehension (of single and multiple documents) as reading activities for multiple texts.

However, multiple-text comprehension research has paid less attention to texts' differences in their content and how they might influence

comprehension and evaluation. Online argumentative texts have a great amount of heterogeneity in the way their content communicates and supports their claims, namely in their *argument types*. In many controversial topics and in various domains, texts may support their claims through *science-based* or through *personal case-based* types of arguments. Recently there have been a few studies taking into account differences similar to the above-mentioned argument types. These studies focused on persuasion (Betsch et al., 2011) and evaluation (Kiili et al., 2022) yielding, however, mixed results. Personal texts about vaccinations have been found to exert more influence on persuasion and decision making in comparison to texts with more scientific content (Betsch et al., 2011). In contrast, Kiili et al. (2022) found texts with more personal content regarding the effect of sugar consumption to be evaluated as less credible and of lower quality of evidence in comparison to other texts presenting more scientific and objective content.

Although comprehension of text and argument must necessarily be a prerequisite for evaluation (Diakidoy et al., 2017; Wolfe & Williams, 2017), a similar interest in the effects of argument type on comprehension and trustworthiness evaluation in a multiple conflicting text context and its possible interactions with other factors has not been documented. Therefore, in this study, we investigated the potential effect of two argument types (science based and personal case based) in single- and multiple-text comprehension and trustworthiness evaluations for texts dealing with vegan nutrition. Information about health and nutrition generates a lot of interest (e.g., on social media platforms), is often conflicting (e.g., Mediterranean vs. keto diet) and, in some cases, its validity is questionable. Nevertheless, readers' prior beliefs and reasons for accessing this information (reading goals) may determine the extent to which information is accepted as valid and trustworthy. Therefore, in this study we measured prior beliefs and manipulated reading goals to examine their influence on single- and multiple-text comprehension and trustworthiness evaluations. Finally, we also measured prior topic knowledge and working memory capacity as they may interfere with comprehension and memory for texts.

## Theoretical Background

Empirical research with multiple documents has focused on variants of expository texts. The Construction-Integration model (Kintsch, 1988, 1998) is the dominant comprehension model focusing on the processes and outcomes of reading primarily single expository texts. Its main principle is that reading consists of a meaning construction phase, aided by the activation of prior knowledge, and an integration phase contributing to the construction of a coherent and elaborated mental

representation of the text. The reading outcome, however, can vary ranging from shallow comprehension of text (surface code) or deeper comprehension (text-based level, situation model level), depending on the degree of integration with prior knowledge structures. Empirical evidence showed, especially, the importance of prior knowledge structures (e.g., high amount of prior knowledge, accurate knowledge, consistent beliefs, etc.) for generating inferences and for achieving deep comprehension of a text (e.g., Kendeou & van den Broek, 2005; Maier & Richter, 2013).

The Documents' Model (Britt et al., 1999; Perfetti et al., 1999) expands upon Kintsch's model by introducing two additional representational structures aimed at evaluating (Intertext Model) and understanding multiple texts (Situations Model). Namely, readers need to comprehend not only the texts separately (single-text comprehension) but also the connections of ideas coming from different texts and, thus, form a coherent representation of the issue discussed across texts (multiple-text comprehension). Second, readers need additionally to evaluate the texts in terms of their trustworthiness by taking into account source information, for example, the expertise of the author (Bråten, Britt, et al., 2011; Britt et al., 1999; Perfetti et al., 1999).

A recent theoretical model (RESOLV model) has emphasized the importance of reading goals in shaping the way individuals engage with texts (Rouet et al., 2017). Specifically, according to the RESOLV model, readers' engagement with the sources' features is also influenced by their reading goals (Britt et al., 2018). Nonetheless, in informal reading situations, readers may not consistently attend to source information, focusing instead on texts' content (e.g., Bråten, Strømsø, & Salmerón, 2011). If this is the case, then it remains an open question as to the characteristics of text content that form the basis for readers' evaluations and whether and the extent to which they interact with reading goals.

## Argumentative Texts and Argument Types

Argumentative texts function to persuade their readers by presenting reasons in support of the claims they promote (Glassner, 2017; Murphy & Alexander, 2004). Argumentative texts differ not only in terms of their structure (e.g., Andiliou et al., 2012) but also in the type of arguments they include. Prior work has focused on rather highly specified distinctions between argument types. For example, Scharrer et al. (2012) investigated the differential effect of causal-based and evidence-based types in judging arguments' strength. Causal-based arguments, which provided causal explanations to claim a connection between biological phenomena, were rated as stronger than evidence-based arguments that relied on statistical data to support such connections. This dichotomy, however, while useful

for analytical purposes, oversimplifies the complexity of actual scientific discourse. In reality, the nuanced nature of scientific writing frequently involves a fusion of both causal explanations and statistical evidence within a single argument or text (Glassner, 2017).

Furthermore, the advent of online texts introduces variability at a more general level, as digital platforms accommodate a wider range of argumentative styles and interdisciplinary approaches that transcend traditional academic boundaries. For example, in the domain of health and nutrition, Google will return texts that either support their claims by relying on scientific knowledge or by presenting personal experience in a storytelling approach. Nevertheless, a notable gap in existing research is evident, with the effects of personal narratives in reading outcomes remaining unexplored. Therefore, we sought to investigate the differential effect of two distinct argument types, namely science based and personal case based. Specifically, we operationalized argument type as arguments supporting a claim either through scientific knowledge (science-based argument type) or through personal case-based experiences (personal case-based argument type).

Science-based argumentative texts rely on scientific information to support a claim, which makes them similar to expository scientific texts in terms of content and structure. A long line of research has shown that the comprehension of scientific expository texts is a challenge (Graesser et al., 2002). A key factor is the lower levels of relevant topic and domain-specific prior knowledge (e.g., Follmer & Sperling, 2019; Graesser et al., 2002; McNamara et al., 1996; Wolfe & Mienko, 2007; Wolfe & Woodwyk, 2010), which is needed for the construction of inferences that increase the coherence and the elaboration of the texts' mental representation (Kraal et al., 2017). It is reasonable to expect that the comprehension of science-based argumentative texts represents a similarly difficult task, since they present scientific information that is, most possibly, unfamiliar to non-expert readers.

However, science-based argumentative texts should be associated with higher trustworthiness ratings in comparison to other argument types, since their evidence-based arguments are of higher quality. Argument quality is directly related to the texts' content such as argument type (e.g., statistical data) (Kuhn, 1991), internal consistency, accuracy, and completeness of information (Shaw, 1996), and lack of any argument fallacies (e.g., Klaczynski et al., 1997). Science-based texts are more likely to have these content-related characteristics in comparison to other texts.

In contrast, personal case-based texts have similarities with narratives, since they introduce characters who have goals, embark on actions, and experience their consequences (Wolfe & Mienko, 2007). Narrative texts have been linked to higher comprehension in comparison to

expository texts. A recent meta-analysis (Mar et al., 2021) showed that narrative texts had better comprehension and memory than expository-like texts. This effect is possibly due to their concreteness (Sadoski et al., 2000) and closeness to readers' prior world knowledge and experiences (e.g., Diakidoy et al., 2005; Wolfe & Woodwyk, 2010). Sadoski et al. (2000) found higher gist recall for texts (persuasive, history narrative and expository) that included real-life experiences and routines, when compared to the more abstract versions of the same text types.

In addition, narratives are more likely to induce mental imagery (e.g., Gerrig, 1993), which has been empirically linked to various emotion experiences during reading (Winkler et al., 2022). Winkler et al. (2022) showed that transportation into a story about an elderly couple's suicide was positively related to the number and intensity of self-reported emotional shifts, that is, changes in dominant emotions like anger, happiness, etc. Transportation and emotions experienced during reading have been shown to contribute further to reading comprehension (Gernsbacher et al., 1992; Gerrig, 1993; Pickren et al., 2022). For example, Pickren et al. (2022) found that texts' emotional arousal at the word-level was positively related to comprehension for various types of texts over and above the influence of readers' and texts' characteristics. Therefore, it is plausible to expect that, due to concreteness and content familiarity and/or the ability to evoke transportation and emotion, the comprehension of the personal case-based argumentative texts to be higher than that of the science-based texts.

Evaluation-wise, personal case-based texts should be evaluated as less trustworthy than science-based texts, since they tend to support claims using a limited number of experiences, personal perspectives, and emotions. Personal opinions and experiences are regarded as weaker arguments in comparison to other more abstract reasons (Means & Voss, 1996) and of lower quality as they are more likely to be associated with argument fallacies, namely, hasty generalization and violations of the law of large numbers (Diakidoy et al., 2017; Glassner, 2017; Klaczynski et al., 1997). However, recognizing personal case-based arguments as of lower quality and, therefore, less trustworthy, is possibly an ideal outcome. Empirical studies showed that adult readers do not routinely take into account argument quality criteria, such as argument fallacies when making evaluation judgments (e.g., Diakidoy et al., 2017). Moreover, the possible emotionality of the personal case-based texts might hinder critical evaluation of their content. For example, Eliades et al. (2012) found lower performance on critical thinking tasks that involved emotional content when compared to those with neutral content, while Mason et al. (2017) showed participants with lower emotional reactivity (low arousal and positive affect) to include more source references in their integrative essays after reading multiple conflicting documents

than participants with higher emotional reactivity. Sourcing, which involves attention to source information, such as authorship and source type, is possibly important for making evaluative judgments for texts (e.g., List & Alexander, 2017). Therefore, it is not certain that readers will take into account the lower argument quality of personal case-based texts when evaluating them.

Nevertheless, in a recent study by Kiili et al. (2022), upper secondary students were able to recognize that a blog post based on a person's observations about the effects of sugar consumption was less credible on the basis of author expertise and evidence quality in comparison to other texts, such as a scientist's research paper or a newspaper article. However, they were not successful with respect to providing relevant and clear justifications for the above-mentioned criteria, namely author expertise and evidence quality. Thus, it is possible that although readers may not possess sufficient knowledge regarding argument quality criteria, their evaluations were aided by the available author and source information. However, in a naturalistic internet search context, source information may be absent or even misleading (e.g., unsupported claims of expertise), or readers may not be motivated to attend to and consider source credibility. Nonetheless, even in the absence of source information, the multiple conflicting text environment may serve to highlight differences in content. These differences may be more prominently represented mentally as readers attempt to resolve the discrepancies (Braasch et al., 2012; Kammerer et al., 2016). Therefore, it is possible that readers might notice and use differences pertaining to argument quality, such as argument type, in favor of science-based texts over personal case-based texts.

## **Belief Consistency**

The extent to which readers notice and consider differences in argument type and, thereby, quality, may also be influenced by their prior beliefs. Research on critical thinking has shown that people are more likely to accept and evaluate positively fallacious arguments that are consistent with their prior beliefs and aspirations in comparison to both high-quality and fallacious arguments that happen to be inconsistent with their beliefs (e.g., Klaczynski et al., 1997). In a multiple conflicting text environment, readers' beliefs are likely to contradict with the claims promoted by some of the texts. Prior work using multiple claim-conflicting texts with expository content has shown a belief consistency effect in both comprehension and evaluation. Specifically, findings have indicated that belief-consistent texts are better comprehended both separately (Abendroth & Richter, 2020; Karimi & Richter, 2021; Maier & Richter, 2013) and combined (Braasch et al., 2022; Maier et al., 2018) in comparison to belief-inconsistent texts when they are part of a conflicting set

of texts. Belief-consistent texts are also regarded by readers as more trustworthy than belief-inconsistent texts (Kobayashi, 2014; van Strien et al., 2016).

Although previous studies have identified additional factors that may influence the belief consistency effect when individuals read multiple conflicting texts (e.g., epistemic beliefs in Karimi & Richter, 2021; the order of texts in Maier & Richter, 2013 etc.), these did not vary the type of the texts (e.g., Maier & Richter, 2013, exclusively used scientific texts) or did not address differences related to text content (e.g., van Strien et al., 2016). The omission of text content variation could limit the generalizability of the findings regarding the belief consistency effect. Consequently, the conclusions drawn may not adequately reflect the complex ways in which differences in text content can influence the relationship between belief consistency and the outcomes of reading multiple texts. Thus, the present study examined the belief consistency effect across varying types of arguments, specifically comparing the impact of science-based texts to that of personal case-based texts.

Moreover, it is important to consider that the narrative content in personal case-based texts may play a significant role in changing opinions and influencing evaluations of the texts read (e.g., Green & Brock, 2000). Transportation into a text's story line and characters was found to be generated regardless of participants' prior attitudes after reading four single texts with different arguments embedded in stories varying in narrativity (e.g., arguments about the importance of salads, tuition fees, etc.) (Schreiner et al., 2018). Transportation has been also found to be positively related to text-consistent stances after reading (Mazzocco et al., 2010). Specifically, Mazzocco et al. (2010) examined in two studies how transportation would affect attitude changes after reading a single narrative text. Their results indicated that those more easily transported into the story were more likely to engage in attitude change in order to align with the text's argument, and that this transportation effect on reader attitudes was not due to other text factors (i.e., rhetorical means) (Mazzocco et al., 2010).

These findings suggest that belief consistency may interact with argument type to reduce or neutralize belief-consistency effects on reading outcomes with personal case-based texts when compared to science-based texts. Therefore, the well-documented belief consistency effect in comprehension and evaluation of texts might not be comparable across different types of text content, such as those that are science based versus those that are personal case based, which possibly stimulate emotional and imagery engagement with their story-telling content.

## **Reading Goals**

People read multiple texts in a variety of formal and informal contexts, and for a variety of reasons: to learn,

to form opinions, and to make important decisions (Betsch et al., 2011). Reasons for reading multiple conflicting texts may influence to some extent whether readers notice and consider differences in argument type or even readers' vulnerability to belief bias. For example, one would expect readers to pay closer attention to the argument and juxtapose claims to their own beliefs when they read in order to make an important decision as opposed to when they read to simply gain information. Many studies have investigated the influence of reading goals in single-text comprehension by manipulating pre-reading instructions (e.g., Diakidoy et al., 2017; van den Broek et al., 2001). In general, findings indicate that reading goals influence both memory for text and inferencing. For example, van den Broek et al. (2001) showed that college students who read a text for study purposes generated more online inferences that helped to construct a coherent mental representation and were more successful in recalling the text offline. Furthermore, Diakidoy et al. (2017) showed that an evaluation goal benefited argumentative text recall to a greater extent than a comprehension goal as participants who read for comprehension were more likely to confuse the text's main claim with some of the arguments presented to support it.

In contrast, multiple-text comprehension research has focused more on post-reading tasks as an indirect manipulation of reading goals (e.g., Wiley et al., 2009). These studies showed that argumentative essay tasks were more effective in helping readers connect information from different texts and maintain a fair balance in representing their content, when compared to tasks designed to gauge learning and understanding, such as summary and essay writing (Bråten & Strømsø, 2009; Le Bigot & Rouet, 2007; Maier & Richter, 2016; Stadler et al., 2014; Wiley et al., 2009).

Setting goals prior to reading is especially important when encountering multiple texts as it can induce goal-relevant processing determining, in part, the contents of the resulting mental model of the text(s) that readers form (Britt et al., 2014; McCrudden, 2018; Rouet & Britt, 2011; van den Broek & Helder, 2017). Recent conceptualizations for reading multiple texts have explicitly incorporated reading goals in their theoretical models (e.g., MD-TRACE model in Rouet & Britt, 2011; RESOLV in Rouet et al., 2017). The RESOLV model (Rouet et al., 2017), specifically, views reading as relevance based and distinguishes context-derived reading goals (*context* model) from the reading goals generated by the reader (*task* model). Even though empirical evidence is lacking, recent studies showed that readers form their own representations for given reading goals and, specifically, tasks. In a recent study, List et al. (2019) examined participants' perceptions for tasks' requirements and showed that almost half of them perceived the

argument task as taking and justifying a position, whereas only a quarter considered it to entail presentation of both sides along with evaluation. Elaboration and source use were less mentioned as requirements for the argument task compared to the essay and the research task, respectively (List et al., 2019). Nonetheless, empirical evidence about the contribution of the task model on actual performance is ambiguous. For example, in a multiple-text context, those who consider argument construction to require citation are more likely to include citations in their responses (List et al., 2019). However, understanding that the task demands entail multiple-document activities (such as detection of similarities and differences between documents and comparison of their content, evaluation of the documents' sources, etc.), did not show any relationship with participants' comprehension of multiple texts or their actual engagement in multiple-document activities online (Schoor et al., 2021).

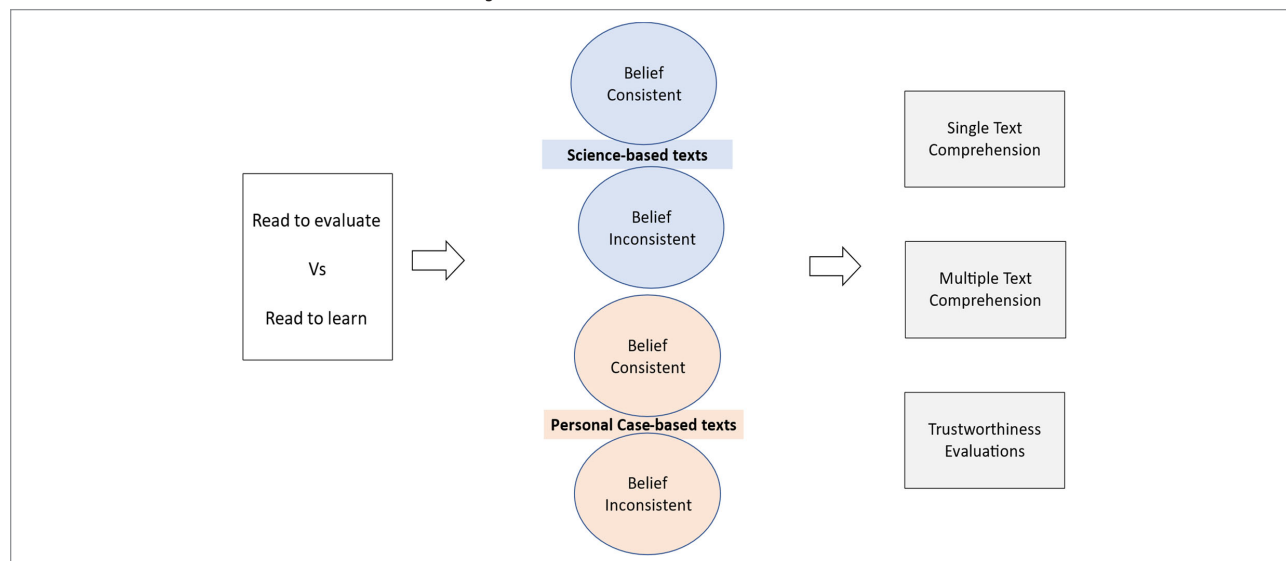
Therefore, in this study, we manipulated directly context reading goals prior to reading in order to examine their effects on single- and multiple-text comprehension and trustworthiness evaluation. We varied reading goals in a way that they would orient readers either toward learning or critical evaluation. We selected these two reading goals because we considered them to be more in line with naturalistic reasons for reading multiple texts in everyday life. They are also possibly more analogous to the recall, summary, and argument writing tasks that have been commonly used in prior research.

## The Present Study

To advance current knowledge, the main purpose of the present study was to examine the possible independent and interactive effects of three critical variables considered together in their relationships, that is, multiple conflicting texts' argument type (science based vs. personal case based), belief consistency (belief consistent vs. belief inconsistent), and reading goals (read to evaluate vs. read to learn) on comprehension (single and multiple) and evaluation (trustworthiness ratings) outcomes. The possible interference of prior topic knowledge and working memory capacity was also controlled for (see Figure 1 for an overview of the study).

Based on prior research (Diakidoy et al., 2005, 2017; Kraal et al., 2017; Maier & Richter, 2016; Sadoski et al., 2000; van Strien et al., 2014, 2016) we hypothesized that argument type, belief consistency, and reading goals would exert significant independent influences on single- and multiple-text comprehension. Specifically, we expected higher comprehension and better memory for personal case-based texts than science-based texts at the single [Hypothesis 1a] and the multiple-text comprehension level [Hypothesis 1b]. We also expected higher

**FIGURE 1**  
**Overview of the Main Variables of the Study**



comprehension and better memory for the belief-consistent than the belief-inconsistent texts at the single [Hypothesis 2a] and the multiple-text comprehension level [Hypothesis 2b]. Similarly, we expected higher comprehension and better memory for participants with an evaluation goal in comparison to a learning goal at the single- and multiple-text comprehension level (Diakidoy et al., 2005, 2017) [Hypotheses 3a and 3b respectively].

On the basis of prior work (Eliades et al., 2012; Escalas, 2004; Kiili et al., 2022), we considered two alternative hypotheses regarding the possible influence of argument type on trustworthiness evaluations. First, it is possible that readers will recognize that the science-based texts have higher argument quality in comparison to the personal case-based texts and, thus, evaluate them as more trustworthy [Hypothesis 4a]. However, in light of findings regarding the negative influence of content with imagery and emotional characteristics on critical evaluation (e.g., Eliades et al., 2012), it is also possible that the personal case-based texts will not be considered less trustworthy than science-based texts [Hypothesis 4b]. In contrast, we expected a more clear influence of prior beliefs on evaluation in favor of the belief-consistent texts [Hypothesis 5] (Kobayashi, 2014).

Despite the lack of empirical evidence regarding the relationship between reading goals and trustworthiness evaluations of multiple conflicting texts, we also considered the possibility of two competing hypotheses. First, we considered it possible that reading to evaluate will result in lower trustworthiness evaluations overall regardless of argument type [Hypothesis 6a] (Diakidoy et al., 2017). However, it is also possible that an interaction would emerge, namely reading to evaluate would support

readers in recognizing the higher quality of science-based arguments resulting in higher trustworthiness ratings for these texts in comparison to the personal case-based texts (Wiley et al., 2009) [Hypothesis 6b].

## Method

### Participants and Procedure

Participants were 95 psychology ( $n = 44$ ) and education ( $n = 51$ ) undergraduates from the University of Cyprus (84.2% females, age  $M = 20.91$  years,  $SD = 2.75$ ). The proposal for the study (aims, materials, and design) was approved by the National Bioethics Committee of Cyprus (Number: CNBC RP 2015.01.47) prior to participant recruitment. Students participated voluntarily, after giving their written consent, and received extra course credit for compensation. Power analysis indicated that the sample size was adequate ( $N \geq 82$ ) for an analysis of variance test with a statistical power of at least .80 ( $\alpha = .05$ ).

The study was divided into two sessions approximately one month apart. Each session lasted 1.5 hours. In Session 1, we administered all pre-reading measures, namely tests for Prior Topic Knowledge, Prior Topic Beliefs, and Working Memory Capacity. For Session 2, participants were randomly divided into the two Reading Goal groups: Read to Learn versus Read to Evaluate. Participants were given both written and oral information about their reading goals. Specifically, participants in the Read to Learn group were instructed to “Read the texts with the aim of learning about vegan nutrition and

its health consequences.” Participants in the Read to Evaluate group were instructed to “Read the texts with the aim of evaluating the arguments presented either in favor or against vegan nutrition.” During Session 2, participants read four experimental texts in a counterbalanced order in order to control for order effects. Then, they rated each text’s trustworthiness and answered three comprehension questions. After reading all the texts, participants completed the multiple-text comprehension task. The text was not available to the participants after the reading was completed. Thus, all post-reading tasks (comprehension and trustworthiness ratings) were completed from memory.

## Experimental Texts

The four experimental texts were presented on separate sheets of paper. They did not contain any source information or other contextual cues, apart from the title and the text itself. Their content was about vegan nutrition and they were harvested from the Internet. The texts were selected to be claim-conflicting and to differ in terms of argument type. Two texts argued against vegan nutrition (con texts) by referring to possible side effects while the other two texts argued in favor of vegan nutrition (pro texts) by focusing on its benefits. Texts promoting the same claim differed in terms of argument type, supporting the claim on the basis of scientific information or on the basis of personal experience. The science-based texts promoted conflicting claims by presenting scientific evidence concerning the health benefits or risks of vegan nutrition. In contrast, the personal case-based texts presented positive or negative effects of vegan nutrition on health, appearance, and/or emotions as experienced by the texts’ protagonists (see Table 1 for examples).

The selected texts were slightly modified to ensure clear differences in argument type and comparable length (687–719 words, 17.9–22.7 words/sentence).

## Pre-Reading Measures

Prior Beliefs were measured with a 10-item Likert-type rating task (1=totally disagree and 6=totally agree) including positive (“Vegan nutrition helps to prevent and fight diseases”) and negative statements (e.g., “Vegan nutrition is a danger to human health”) about vegan nutrition. Negative statements were reverse scored, so that higher total scores would indicate more positive beliefs toward vegan nutrition ( $M=3.77$ ,  $SD=.65$ ). The rating task’s reliability was satisfactory (Cronbach’s  $\alpha=.77$ ), and a median split ( $Mdn=3.80$ ) was used to determine Belief Consistency (pro vs. against vegan nutrition).

Prior Topic Knowledge was measured with a set of 23 statements about nutrition and biology, each followed by a Likert-type scale (1 strongly disagree–6 strongly agree). Some of the statements were accurate ( $n=12$ ) (e.g., “Fructose and lactose are carbohydrates”) and the rest were incorrect (e.g., “Saturated and trans-fat are good fat”). False statements were reverse scored, with higher total scores reflecting higher levels of prior topic knowledge. Participants’ ratings were summed across items and divided by the total number of items in the test, to create a mean score for Prior Topic Knowledge. The test had adequate reliability (Cronbach’s  $\alpha=.70$ ).

Working memory capacity was measured with a reading span test that was constructed based on the test used by Singer et al. (1992). The test used in the present study included 18 sets of sentences (63 sentences in total) that were derived from sources with different textual content (expository and narrative). The number of sentences

**TABLE 1**  
**Examples of Science-Based and Personal Case-Based Arguments in Experimental Texts**

| Texts                        | Example excerpt   |
|------------------------------|---|
| Science-based Con text       | “Good cholesterol helps reduce bad cholesterol (LDL) by transferring it to the liver where it can be reprocessed. It also helps maintaining the internal walls of the blood vessels. Therefore, the heart needs cholesterol and if it is not produced, you will end up with heart failure”  |
| Science-based Pro text       | “Plant-based nutrition has a minimal amount of saturated and trans fats. By contrast, most saturated fats are derived from animal products (...) It appears that saturated and trans fats, like cholesterol, from animal products increase the risk of arterial obstruction and heart disease”  |
| Personal case-based Con text | “During the last six months of her vegan diet, 26-year-old Pamela Stubbart saw her blood sugar levels go up and experienced intense cravings. She could not concentrate on anything else except whether or not she had to eat an egg (...) While initially her energy levels skyrocketed, then they fell sharply and she felt extremely tired. In fact, panic attacks were almost a daily situation”  |
| Personal case-based Pro text | “I also read some informal reports from older vegetarians on how they felt their diet delayed their aging. A friend of my sister became a vegan and, almost overnight, her battle with irritable bowel syndrome (spastic colitis) stopped. Dairy products, it seemed, upset her stomach and, by following a plant-based diet, her problems were eliminated. (...) Also, I almost immediately discovered a whole new community of friends I so long lost, not only vegans, but also those who are interested in better health and environmental sustainability. (...)” |

changed across sets in order to reflect different difficulty levels: three sets had two sentences, six sets had three sentences, six sets had four sentences, and three sets had five sentences. A higher number of sentences in a set reflected a higher difficulty level. Within each difficulty level, the length of sentences also varied: there were 10-word, 12-word, and 14-word sentences. The reading span test was administered through a computer for each participant. Participants had approximately 3–5 minutes to read each set before the cloze and the recall tasks were presented automatically on separate screens. For the cloze task, participants had to fill in two words missing from a randomly selected sentence from the set. For the recall task, participants had to fill in the last word of each sentence presented in the set. A pilot study with a different but comparable sample ( $N=110$ , 78.8% females,  $M_{\text{age}}=20.6$ ,  $SD=4.1$ ) indicated that a strict scoring of the cloze task resulted in low overall performance ( $M=.20$ ,  $SD=.19$ ). Therefore, we decided to apply a more lenient scoring giving credit for correct performance on the recall task when performance in the corresponding cloze task was partially correct (providing at least one of the two words). Participants' points in the recall tasks were summed and divided by the total number of sets in the test to create a score reflecting the proportion of correctly recalled words across sets, and the test's reliability was satisfactory (Cronbach's  $\alpha=.78$ ).

## Post-Reading Measures

Single-Text Comprehension was assessed with three open-ended questions per text. In total, participants answered 12 open-ended questions for all four texts. Each question targeted implicit relations between the text's parts (Rukavina & Daneman, 1996). This was achieved by dividing each text into sentences and sub-topics and then identifying the links (e.g., causal, categorical, etc.) between them. These links were used as the basis for designing the questions and prototype answers for scoring purposes. For example, in order to correctly answer the question asking about the relationship between heart problems and vegan nutrition, participants had to infer the links connecting information about the liver's function, cholesterol, sugar, and artery blockage. Participants' answers for each question were compared against the corresponding prototype answer and received one point for each correctly recalled text idea. Then, we summed all points received across all questions per text and converted them into a proportion score based on the maximum score one could obtain for each text. Apart from the main rater (one of the authors), another independent rater scored 30% of participants' answers for inter-rater reliability purposes. Inter-rater agreement was 83.3–100% for single-text comprehension answers per text,  $K=.66$  to  $K=1$ , all  $p<.001$ . Differences were resolved in conference.

Multiple-Text Comprehension was assessed with a single synthetic question asking participants to delineate the advantages and disadvantages of vegan nutrition based on all texts in the set. The selection of this particular question format aimed to encourage the synthesis of information from the various texts in the set. This task was chosen in order to minimize biased responses that stem from pre-existing beliefs (e.g., van Strien et al., 2014, 2016), which might be included in an essay task that is commonly used to assess comprehension of multiple texts (e.g., Anmarkrud et al., 2013). We used the texts' sentences and the previously identified links between them in order to construct a prototype answer to the Multiple-Text Comprehension question. The prototype answer included 81 text-based ideas from all texts (14 to 28 text-based ideas for each text). Participants' answers were compared against the corresponding prototype answer and received one point for each correctly recalled text idea. Then, we summed all points received and converted them into a proportion score based on the maximum score one could obtain from the Multiple-Text Comprehension question. Additionally, we constructed four text-specific proportion scores, reflecting the proportion of correctly recalled ideas from each text. A second independent rater scored 30% of participants' answers for inter-rater reliability purposes. Inter-rater agreement was 83.3–90%,  $K=.71$ ,  $p<.001$ . Differences were resolved in conference.

The difference between the Single- and the Multiple-Text Comprehension measures lies in the different kinds of processing required to complete them. The Single-Text Comprehension task targeted inferring implicit connections and integrating information contained within a single document. In contrast, the Multiple-Text Comprehension task required participants to recall arguments in favor and against vegan nutrition presented across texts. To complete the task participants needed to recall comparable and contrasting information from different texts and perspectives. The integration of different perspectives and the representation of links or contrasts between documents is an integral part of the Documents' Model (Britt et al., 1999; Perfetti et al., 1999).

We measured Trustworthiness evaluations with a single Likert-type rating item (1 = totally disagree and 6 = totally agree) per text, following the approach adopted in previous research (e.g., Bråten, Strømsø, & Salmerón, 2011b). Specifically, we asked participants to rate the level of their agreement to the following statement: "I consider the text as credible and I can trust the content presented."

## Results

Preliminary analyses indicated that all dependent measures and covariates were normally distributed



(skewness <1) with homogeneous variances across Reading Goal groups ( $p > .10$ ). Moreover, Reading Goal groups' performance was comparable across covariates and comprehension measures ( $p > .05$ ) except for Trustworthiness Evaluations ( $p < .01$ ). Main analyses followed the logic of Mixed Analysis of Covariance with Argument Type (science based vs. personal case based) and Prior Belief (belief consistent vs. belief inconsistent) as the within-subject factors, Reading Goals as the between-subjects variable and Prior Topic Knowledge and Working Memory Capacity as the covariates. Table 2 shows descriptive statistics and correlations for main dependent measures and covariates.

### Comprehension

Contrary to expectations (Hypotheses 1a, 2a, and 3a), the results for Single-Text Comprehension indicated no significant effects of Argument Type ( $F(1, 89) = 3.36, p = .070, \eta^2_p = .04$ ), Belief Consistency ( $F(1, 89) = 2.53, p = .115, \eta^2_p = .03$ ), and Reading Goals ( $F(1, 89) = 2.80, p = .098, \eta^2_p = .03$ ). Nevertheless, in line with Hypothesis 1a, it can be seen from Table 3 that, overall, personal case-based texts ( $M = .30, SD = .10$ ) were better comprehended than science-based texts ( $M = .25, SD = .09$ ), (paired  $t = 4.94, p < .001$ , Cohen's  $d = 0.52$ ) regardless of belief consistency ( $p > .05$ ). Finally, Prior Topic Knowledge ( $F(1, 89) = 16.88, p < .001, \eta^2_p = .16$ ) and Working Memory Capacity ( $F(1, 89) = 5.33, p = .023, \eta^2_p = .06$ ) positively contributed to Single-Text Comprehension outcomes (see also Table 2).

Regarding Multiple Text Comprehension and in line with Hypothesis 1b, there was a significant effect of Argument Type,  $F(1, 89) = 5.46, p = .022, \eta^2_p = .06$ . Participants recalled more information from personal case-based texts ( $M = .22, SD = .08$ ) compared to the science-based texts ( $M = .15, SD = .08$ ),  $t = -8.57, p < .001$ , Cohen's  $d = 0.87$ . Although the main effect of the Reading Goal was not significant ( $p > .05$ ), there was a

significant two-way interaction between Reading Goal and Belief Consistency ( $F(1, 89) = 7.24, p = .009, \eta^2_p = .08$ ) and a three-way interaction between Reading Goals, Prior Belief Consistency, and Argument type ( $F(1, 89) = 8.62, p = .004, \eta^2_p = .09$ ). It can be seen from Table 4 and Figure 2 that participants recalled more information from the belief-consistent personal case-based texts when reading to learn (paired  $t = 1.91, p < .05$ , Cohen's  $d = 0.46$ ) and more information from the belief-inconsistent personal case-based texts when reading to evaluate (paired  $t = -2.69, p < .01$ , Cohen's  $d = 0.60$ ). In contrast, performance for the belief-consistent and belief-inconsistent science-based texts was comparable across Reading Goal groups ( $p > .05$ ). Finally, Prior Topic Knowledge ( $F(1, 89) = 7.12, p = .009, \eta^2_p = .07$ ) and Working Memory Capacity ( $F(1, 89) = 11.44, p = .001, \eta^2_p = .11$ ) contributed positively to Multiple-Text Comprehension outcomes (see also

**TABLE 3**  
Single-Text Comprehension as a Function of Reading Goals, Belief Consistency, and Argument Type

| Single-text comprehension | Read to learn |           | Read to evaluate |           | Cohen's <i>d</i> |
|---------------------------|---------------|-----------|------------------|-----------|------------------|
|                           | <i>M</i>      | <i>SD</i> | <i>M</i>         | <i>SD</i> |                  |
| Belief Consistent         |               |           |                  |           |                  |
| Science based             | .23           | .14       | .26              | .12       | 0.23             |
| Personal case based       | .30           | .14       | .29              | .14       | 0.07             |
| Cohen's <i>d</i>          | 0.50          |           | 0.23             |           |                  |
| Belief Inconsistent       |               |           |                  |           |                  |
| Science based             | .23           | .10       | .25              | .10       | 0.20             |
| Personal case based       | .27           | .13       | .34              | .14       | 0.52             |
| Cohen's <i>d</i>          | 0.34          |           | 0.74             |           |                  |

**TABLE 2**  
Descriptive Statistics and Correlations for Pre-Reading and Post-Reading Measures

| Variable   | <i>M</i> | <i>SD</i> | 1     | 2     | 3     | 4     | 5    | 6 |
|--|----------|-----------|-------|-------|-------|-------|------|---|
| 1. Prior beliefs                                     | 3.77     | .65       | -     |       |       |       |      |   |
| 2. Prior Topic Knowledge                             | 4.27     | .48       | -.20* | -     |       |       |      |   |
| 3. Working Memory Capacity                           | .25      | .13       | -.09  | .20   | -     |       |      |   |
| 4. Single-Text Comprehension (overall mean score)    | .27      | .08       | -.03  | .43** | .32** | -     |      |   |
| 5. Multiple-Text Comprehension (overall mean score)  | .18      | .07       | -.01  | .32** | .38** | .57** | -    |   |
| 6. Trustworthiness Evaluations (overall mean rating) | 3.73     | 0.84      | .25*  | .02   | -.17  | -.24* | -.08 | - |

\* $p < .05$ . \*\* $p < .01$ .

**TABLE 4**  
Multiple-Text Comprehension as a Function of Reading Goals, Belief Consistency, and Argument Type

| Multiple-text comprehension | Read to learn |           | Read to evaluate |           | Cohen's <i>d</i> |
|-----------------------------|---------------|-----------|------------------|-----------|------------------|
|                             | <i>M</i>      | <i>SD</i> | <i>M</i>         | <i>SD</i> |                  |
| Belief Consistent           |               |           |                  |           |                  |
| Science based               | .15           | .09       | .16              | .09       | 0.11             |
| Personal case based         | .25           | .11       | .20              | .10       | 0.48             |
| Cohen's <i>d</i>            | 1.00          |           | 0.42             |           |                  |
| Belief Inconsistent         |               |           |                  |           |                  |
| Science based               | .14           | .10       | .15              | .11       | 0.10             |
| Personal case based         | .18           | .11       | .24              | .10       | 0.57             |
| Cohen's <i>d</i>            | 0.38          |           | 0.86             |           |                  |

**TABLE 5**  
Trustworthiness Evaluations as a Function of Reading Goals, Belief Consistency, and Argument Type

| Trustworthiness evaluations | Read to learn |           | Read to evaluate |           | Cohen's <i>d</i> |
|-----------------------------|---------------|-----------|------------------|-----------|------------------|
|                             | <i>M</i>      | <i>SD</i> | <i>M</i>         | <i>SD</i> |                  |
| Belief Consistent           |               |           |                  |           |                  |
| Science based               | 4.20          | 0.99      | 3.75             | 1.03      | 0.45             |
| Personal case based         | 4.08          | 1.18      | 3.60             | 1.17      | 0.41             |
| Cohen's <i>d</i>            | 0.11          |           | 0.14             |           |                  |
| Belief Inconsistent         |               |           |                  |           |                  |
| Science based               | 3.97          | 1.12      | 3.47             | 0.97      | 0.48             |
| Personal case based         | 3.78          | 1.16      | 3.02             | 1.23      | 0.64             |
| Cohen's <i>d</i>            | 0.17          |           | 0.41             |           |                  |

Table 2). No other main effects were found, thus, Hypotheses 2b and 3b were not supported.

### Trustworthiness

Analyses for Trustworthiness Evaluations indicated a significant main effect for Reading Goals only,  $F(1, 89) = 9.64$ ,  $p = .003$ ,  $\eta^2_p = .10$ . Participants who read to evaluate gave lower evaluations overall ( $M = 3.51$ ,  $SD = 0.79$ ) in comparison to the evaluations of those who read to learn ( $M = 4.01$ ,  $SD = 0.83$ ,  $t = 3.00$ ,  $p = .003$ , Cohen's  $d = 0.62$ ). Nevertheless, in line with Hypothesis 4a, it can be seen from Table 5 that, overall, science-based texts were evaluated as more trustworthy than personal case-based texts ( $M = 3.85$ ,  $SD = 0.92$  and  $M = 3.61$ ,  $SD = 0.99$ , paired  $t = 2.58$ ,  $p = .011$ , Cohen's  $d = 0.25$ ). Similarly, belief-consistent texts were evaluated as more trustworthy than belief-inconsistent texts ( $M = 3.91$ ,

$SD = 0.97$  and  $M = 3.55$ ,  $SD = 1.00$ , respectively, paired  $t = 3.38$ ,  $p = .001$ , Cohen's  $d = 0.37$ ).

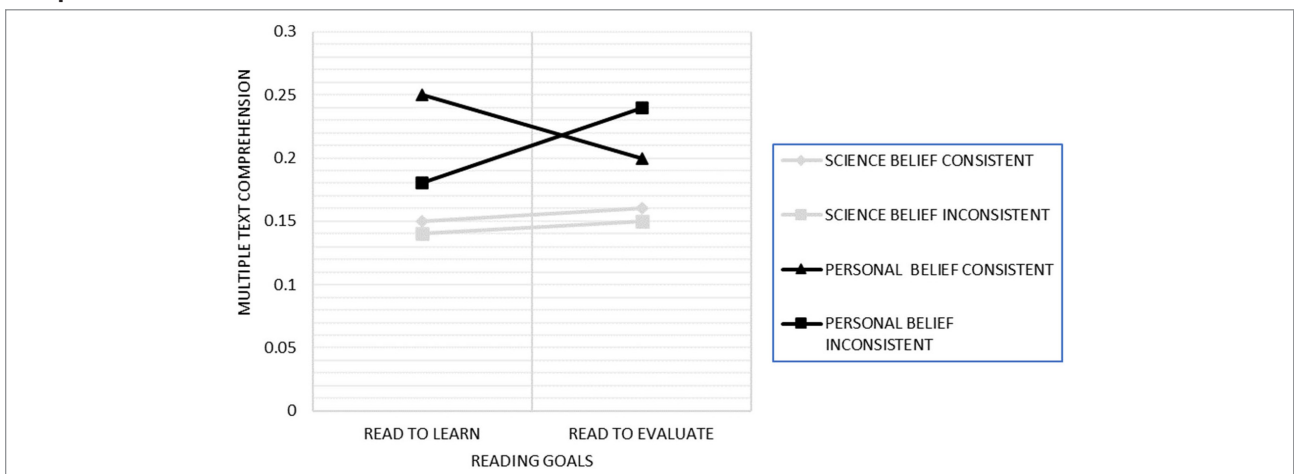
### Discussion

The study examined the effects of argument type, belief consistency, and reading goals in multiple conflicting text comprehension and evaluation. The findings highlighted the influence of argument type on multiple-text comprehension and that of reading goals on trustworthiness evaluations.

### Comprehension

The findings provided support for our hypothesis concerning the influence of argument type on multiple-text

**FIGURE 2**  
Three-Way Interaction between Argument Type, Belief Consistency, and Reading Goals on Multiple-Text Comprehension



comprehension (Hypothesis 1b), but only partial support for its influence at the single-text comprehension level (Hypothesis 1a). In general, texts that capitalize on personal experiences and emotions to support their claims were better comprehended and their content was more memorable than texts that capitalize on scientific information. These findings are consistent with studies showing single-text comprehension and memory advantage for narrative texts with concrete content in comparison to expository texts with more abstract content (Diakidoy et al., 2005; Graesser et al., 2002; Kraal et al., 2017; McNamara et al., 2011; Sadoski et al., 2000). The present study documents this advantage at the multiple-text comprehension level as well.

Contrary to our expectations (Hypotheses 2a, 2b, 3a, and 3b), neither belief consistency nor reading goals had an independent influence on single- and multiple-text comprehension. With respect to belief consistency, it is possible that the overall moderate prior beliefs of our participants, the counter-balanced order of text presentation, and the recall task that is less likely to elicit personal opinions in comparison to an essay task may have served to reduce the manifestation of belief bias in comprehension and recall (Hart et al., 2009; Maier & Richter, 2013; van Strien et al., 2016; Wiley et al., 2018). However, the significant three-way interaction also suggests that argument type may have masked the effects of belief consistency and reading goal. Participants who read to learn recalled more information from belief-consistent texts, showing thus a belief bias effect. This effect has been documented in many multiple-text comprehension studies (e.g., Kobayashi, 2014; Maier et al., 2018; van Strien et al., 2014, 2016). On the contrary, those who read to evaluate recalled more information from the belief-inconsistent texts. Critical thinking research has shown that people are more likely to scrutinize and question arguments and evidence inconsistent with their own beliefs (Klaczynski et al., 1997). Therefore this finding may be the side-effect of the deeper processing associated with motivated reasoning in the context of evaluation tasks.

These results plausibly reflect different text processing due to reading goals. Namely, participants possibly understood that the read to learn goal entails memory-oriented processing which leaves open the possibility of belief bias manifesting in comprehension and recall tasks. In contrast, participants who read to evaluate might have engaged more in validation and evaluation processes resulting in the reverse effect (Diakidoy et al., 2017; Maier & Richter, 2016). However, it is interesting to note that these bias effects were manifested only in connection with the personal case-based texts. It is possible that the objective nature of reporting facts and processes in conjunction with the higher unfamiliarity and difficulty of the science-based texts (as indicated by the lower levels of recall), left little room for knowledge- and belief-based evaluations

and contributed to more memory-oriented processing regardless of reading goals. Nevertheless, the present results, taken together, are partly in line with prior work suggesting that reading goals that orient readers toward more critical processing contribute to more accurate and balanced recall of argumentative texts (Diakidoy et al., 2017; Maier & Richter, 2016).

With respect to the covariates, prior topic knowledge and working memory capacity significantly and positively influenced both single- and multiple-text comprehension outcomes. Prior topic knowledge has been widely documented as an important factor in aiding single- and multiple-text comprehension (e.g., Bråten et al., 2014). In addition, reading span tasks as a measure of working memory capacity have been found to significantly influence single (e.g., Nouwens et al., 2016) and multiple-text comprehension measures (Banas & Sanchez, 2012; Braasch et al., 2014). Moreover, working memory capacity can be expected to play a role in the comprehension of argumentative texts, as readers need to keep important argument elements (claims and supporting reasons, at least) and their interrelations active in working memory (e.g., Diakidoy et al., 2017). Finally, the comprehension measures used in this study may have highlighted the contribution of both prior knowledge and working memory capacity, considering that they required inferring connections between text parts at the single-text level and maintaining information associated with opposing claims and arguments at the multiple-text level (Baddeley, 2012; Kraal et al., 2017).

Moreover, our comprehension measures—single-text comprehension through bridging questions and multiple-text comprehension through the recall of arguments for and against vegan nutrition—mirror the complex cognitive processes individuals engage in daily while navigating online information. The increasing reliance on the Internet as the dominant source for educating ourselves on health and nutrition necessitates a nuanced understanding of how we comprehend and synthesize information from texts with diverse argument types. Particularly in the context of health and nutrition, where personal anecdotes can be quite influential in opinion formation and decision making (e.g. Betsch et al., 2011), our argument type as the main independent variable illuminates the differential impacts the argument types may have on comprehension and the subsequent integration of information.

## Evaluation

Consistent with our hypothesis (6a) and prior research (Diakidoy et al., 2017), participants who read to evaluate were stricter in their trustworthiness ratings than those who read to learn. However, it is noteworthy that the

reading goal was the only factor influencing trustworthiness evaluations in this study. Argument type and belief consistency did not have a significant independent or interactive influence on trustworthiness ratings (Hypotheses 4a, 5, and 6b). These findings are in contrast with empirical evidence demonstrating more positive evaluations for belief consistent over belief inconsistent content (Kobayashi, 2014; McCrudden & Barnes, 2016). Nevertheless, our participants tended to consider science-based and belief-consistent texts as more trustworthy than personal case-based and belief-inconsistent texts. These results suggest that participants did take into consideration to some extent argument or possible source quality (Glassner, 2017; Kiili et al., 2022; List & Alexander, 2017) and their own beliefs. Therefore, it is possible that the influence of reading goals may have served to mask in part the effects of both argument type and belief consistency.

### **Limitations and Future Directions**

In this study, we manipulated argument type on the basis of whether a text capitalized on scientific evidence or personal experience and emotion to promote a claim. This distinction is prominent in naturally occurring Internet texts focusing on health and nutrition, it can be associated with informal reasoning fallacies such as hasty generalization and false cause (e.g., Ricco, 2007), and it can subsume more specific distinctions at a subordinate level (e.g., Scharrer et al., 2012). However, it remains an open question whether the distinction can generalize across topics, beyond health and nutrition. Moreover, it must be noted that our argument-driven text selection did not take into account the valence of text content. The texts promoting a claim in favor of vegan nutrition focused on the positive aspects and consequences of vegan nutrition, whereas the texts promoting a claim against vegan nutrition focused on the negative aspects and consequences. However, there is some evidence that content discussing dangers for human health may be considered more trustworthy than positive information (e.g., Siegrist & Cvetkovich, 2001). Therefore, future work needs to explore this possibility in the context of multiple-text comprehension and evaluation.

Our findings are, at least in part, consistent with those of other studies that have examined the influence of narrative content on comprehension (e.g., Sadoski et al., 2000), belief bias in recall (e.g., Maier & Richter, 2013), and reading goals on texts' evaluation (e.g., Diakidoy et al., 2017). This consistency lends a degree of external validity to our results, suggesting that the observed patterns and effects possibly reflect genuine phenomena. However, it must be noted that the observed power associated with the testing of the effects of our three main variables ranged from .61 to .73 in the full model, indicating that the sample size was

somewhat limited given the design of the study. This leaves open the possibility of true effects being masked and limits the generalizability of the findings. Therefore, further research is needed to examine the influence of argument types on single- and multiple-text comprehension and trustworthiness evaluation.

In this study, we measured multiple-text comprehension by asking participants to provide a list of the advantages and disadvantages of vegan nutrition based on the texts read. We chose this recall task as opposed to a synthetic or argumentative essay in order to minimize task interpretation influences and limited use of all textual sources (List et al., 2019; Rouet et al., 2017; van Strien et al., 2014, 2016; Wiley et al., 2014, 2018). It is also possible that the argument task is interpreted differently by readers (Rouet et al., 2017), which, may explain evidence showing that the argumentative essay task influenced comprehension differently based on the level of prior knowledge and epistemic beliefs (Bråten & Strømsø, 2009; Gil et al., 2010). Although our instructions emphasized the use of information from all texts, one could question the extent to which this measure reflects across-text integration. Specifically, one could argue that the task reflects single-text representations as opposed to a single across-text integrated one. However, even if that is the case, to complete the recall task participants had to activate single-text representations, recognize across text claim conflicts (e.g., Britt et al., 1999), and organize their recall accordingly. Nevertheless, future work can be designed to provide for direct comparisons between recall and essay tasks.

Additionally, it is important for future research to address questions regarding the direct and indirect relationships between person, text, and task factors and comprehension and evaluation outcomes. Rouet et al. (2017) suggested that multiple text reading outcomes are predicted by the goals set in the beginning. It is, therefore, pertinent to investigate a model wherein the effects of reading goals on comprehension—whether of single or multiple texts—are mediated by evaluations of text trustworthiness. This approach not only contributes to the existing literature, which has shown a link between evaluation and comprehension in the context of multiple documents (Bråten et al., 2009; Wiley et al., 2009), but also explores the predictive power of reading goals. Moreover, it is important to investigate also a possible model wherein reading goals' effect on comprehension is mediated by argument types and /or belief consistency. This can inform the design of targeted interventions regarding different instructional designs, textual materials, etc. Overall, the investigation of the above path models would offer insights into the possible processes that underlie the observed relationships and contribute to more robust results and educational practices.

## Conclusions and Implications

Although in naturalistic Internet search contexts, source information is typically provided, knowledge about the source or the criteria to evaluate its quality is often lacking (Britt et al., 2014). Moreover, readers may fail to attend to the source information, focusing, instead, and using text content to guide comprehension and trustworthiness evaluations (Bråten et al., 2009). The present study provided evidence that a characteristic associated with text content, such as argument type, can influence the comprehension of multiple conflicting texts and moderate the effects of other person- and task-characteristics, such as prior beliefs and reading goals. Texts that present personal experiences and conform to narrative structures are more memorable than texts presenting scientific evidence in support of claims. The memorability of stories, however, can be a double-edged sword. Although it can be exploited instructionally to support learning (e.g., Wolfe & Mienko, 2007), it can also contribute to, sometimes unwarranted, changes in behaviors, beliefs, and attitudes, as prior social and consumer research has shown (Escalas, 2004; Mazzocco et al., 2010). The present findings suggested that young adult readers may be aware of the lower argument quality associated with personal stories, considering their tendency to give lower trustworthiness evaluations. However, lower trustworthiness, just as lower argument quality, does not necessarily prevent readers from being influenced by the message, especially if that message is close to their own experience, aspirations, and fears (e.g., Klaczynski et al., 1997). Nevertheless, learning and applying context-appropriate criteria for evaluating both arguments and evidence is amenable to instruction and can contribute to the development of a more critical stance toward information.

Our results also have more general but significant implications for health communication strategies, particularly in the design and dissemination of public health messages through texts. Incorporating personal narratives and experiences into texts could enhance the memorability of the message and, in turn, its impact. However, it must also be acknowledged that personal experiences are and cannot be generalized across people and situations. Therefore, this approach also highlights the need to ensure that these narratives accurately represent scientific content to avoid the spread of misinformation.

## Conflict of Interest statement

None.

## Funding Information

The study was funded in part by the A.G. Leventis Foundation.

## REFERENCES

- Abendroth, J., & Richter, T. (2020). Text-belief consistency effect in adolescents' comprehension of multiple documents from the web. *Journal for the Study of Education and Development*, 43(1), 60–100. <https://doi.org/10.1080/02103702.2019.1692289>
- Andiliou, A., Ramsay, C. M., Murphy, M. P., & Fast, J. (2012). Weighing opposing positions: Examining the effects of intratextual persuasive messages on students' knowledge and beliefs. *Contemporary Educational Psychology*, 37(2), 113–127. <https://doi.org/10.1016/j.cedpsych.2011.10.001>
- Anmarkrud, Ø., McCrudden, M., Bråten, I., & Strømsø, H. I. (2013). Task-oriented reading of multiple documents: Online comprehension processes and offline products. *Instructional Science*, 41, 873–894. <https://doi.org/10.1007/s11251-013-9263-8>
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63(1), 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Banas, S., & Sanchez, C. A. (2012). Working memory capacity and learning underlying conceptual relationships across multiple documents. *Applied Cognitive Psychology*, 26, 594–600. <https://doi.org/10.1002/acp.2834>
- Betsch, C., Ulshöfer, C., Renkewitz, F., & Betsch, T. (2011). The influence of narrative v. statistical information on perceiving vaccination risks. *Medical Decision Making*, 31(5), 742–753. <https://doi.org/10.1177/0272989X11400419>
- Braasch, J. L. G., Bråten, I., Strømsø, H. I., & Anmarkrud, Ø. (2014). Incremental theories of intelligence predict multiple document comprehension. *Learning and Individual Differences*, 31, 11–20. <https://doi.org/10.1016/j.lindif.2013.12.012>
- Braasch, J. L. G., Haverkamp, Y. E., Latini, N., Shaw, S., Arshad, M. S., & Bråten, I. (2022). Belief bias when adolescents read to comprehend multiple conflicting texts. *Reading and Writing: An Interdisciplinary Journal*, 35, 1759–1785. <https://doi.org/10.1007/s11145-022-10262-w>
- Braasch, J. L. G., Rouet, J.-F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory and Cognition*, 40, 450–465. <https://doi.org/10.3758/s13421-011-0160-6>
- Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, H. I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learning and Instruction*, 30, 9–24. <https://doi.org/10.1016/j.learninstruc.2013.11.002>
- Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist*, 46(1), 48–70. <https://doi.org/10.1080/00461520.2011.538647>
- Bråten, I., Strømsø, H., & Salmerón, L. (2011). Trust and mistrust when students read multiple information sources about climate change. *Learning and Instruction*, 21(2), 180–192. <https://doi.org/10.1016/j.learninstruc.2010.02.002>
- Bråten, I., & Strømsø, H. I. (2009). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes*, 47, 1–31. <https://doi.org/10.1080/01638530902959646>
- Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly*, 44(1), 6–28. <https://doi.org/10.1598/RRQ.44.1.1>

- Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence* (pp. 209–233). Erlbaum.
- 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(2), 104–122. <https://doi.org/10.1080/00461520.2014.916217>
- Britt, M. A., Rouet, J.-F., & Durik, A. (2018). *Literacy beyond text comprehension: A theory of purposeful reading*. Routledge.
- Coiro, J. (2021). Toward a multifaceted heuristic of digital reading to inform assessment, research, practice, and policy. *Reading Research Quarterly*, 56(1), 9–31. <https://doi.org/10.1002/rrq.302>
- Diakidoy, I. N., Ioannou, M. C., & Christodoulou, S. A. (2017). Reading argumentative texts: Comprehension and evaluation goals and outcomes. *Reading and Writing*, 30, 1869–1890. <https://doi.org/10.1007/s11145-017-9757-x>
- Diakidoy, I. N., Stylianou, P., Karefillidou, C., & Papageorgiou, P. (2005). The relationship between listening and reading comprehension of different types of text at increasing grade levels. *Reading Psychology*, 26(1), 55–80.
- Eliades, M., Mansell, W., Stewart, A. J., & Blanchette, I. (2012). An investigation of belief-bias and logicity in reasoning with emotional contents. *Thinking & Reasoning*, 18, 461–479. <https://doi.org/10.1080/13546783.2012.713317>
- Escalas, J. E. (2004). Imagine yourself in the product: Mental simulation, narrative transportation, and persuasion. *Journal of Advertising*, 33(2), 37–48. <https://doi.org/10.1080/00913367.2004.10639163>
- Follmer, D. J., & Sperling, R. A. (2019). A latent variable analysis of the contribution of executive function to adult readers' comprehension of science text: The roles of vocabulary ability and level of comprehension. *Reading and Writing: An Interdisciplinary Journal*, 32(2), 377–403. <https://doi.org/10.1007/s11145-018-9872-3>
- Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. (1992). Do readers mentally represent characters' emotional states? *Cognition and Emotion*, 6(2), 89–111. <https://doi.org/10.1080/02699939208411061>
- Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*. Yale University 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. <https://doi.org/10.1016/j.cedpsych.2009.11.002>
- Glassner, A. (2017). Evaluating arguments in instruction: Theoretical and practical directions. *Thinking Skills and Creativity*, 24, 95–103. <https://doi.org/10.1016/j.tsc.2017.02.013>
- Graesser, A. C., León, J. A., & Otero, J. (2002). Introduction to the psychology of science text comprehension. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 1–15). Erlbaum.
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701–721. <https://doi.org/10.1037/0022-3514.79.5.701>
- Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, 135(4), 555–588. <https://doi.org/10.1037/a0015701>
- Kammerer, Y., Kalbfell, E., & Gerjets, P. (2016). Is this information source commercially biased? How contradictions between web pages stimulate the consideration of source information. *Discourse Processes*, 53(5–6), 430–456. <https://doi.org/10.1080/0163853X.2016.1169968>
- Karimi, M. N., & Richter, T. (2021). Biased representations of controversial information: Certainty and justification beliefs as moderators. *Contemporary Educational Psychology*, 66, 101995. <https://doi.org/10.1016/j.cedpsych.2021.101995>
- Kendeou, P., & van den Broek, P. (2005). The effects of Readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology*, 97(2), 235–245. <https://doi.org/10.1037/0022-0663.97.2.235>
- Kiili, C., Bråten, I., Strømsø, H. I., Hagerman, M. S., Rääkkönen, E., & Jyrkiäinen, A. (2022). Adolescents' credibility justifications when evaluating online texts. *Education and Information Technologies*, 27(6), 7421–7450. <https://doi.org/10.1007/s10639-022-10907-x>
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–Integration model. *Psychological Review*, 95, 163–182. [https://doi.org/10.1016/S0166-4115\(08\)61551-4](https://doi.org/10.1016/S0166-4115(08)61551-4)
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Klaczynski, P. A., Gordon, D. H., & Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning bias. *Journal of Educational Psychology*, 89, 470–485. <https://doi.org/10.1037/0022-0663.89.3.470>
- Kobayashi, K. (2014). Students' consideration of source information during the reading of multiple texts and its effect on intertextual conflict resolution. *Instructional Science*, 42, 183–205. <https://doi.org/10.1007/s11251-013-9276-3>
- Kraal, A., Koornneef, A. W., Saab, N., & van den Broek, P. W. (2017). Processing of expository and narrative texts by low- and high-comprehending children. *Reading and Writing: An Interdisciplinary Journal*, 31, 2017–2040. <https://doi.org/10.1007/s11145-017-9789-2>
- Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.
- Le Bigot, L., & Rouet, J. F. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research*, 39, 445–470. <https://doi.org/10.1080/10862960701675317>
- List, A., & Alexander, P. A. (2017). Cognitive affective engagement model of multiple source use. *Educational Psychologist*, 52(3), 182–199. <https://doi.org/10.1080/00461520.2017.1329014>
- List, A., Du, H., & Wang, Y. (2019). Understanding students' perceptions of task assignments. *Contemporary Educational Psychology*, 59, 101801. <https://doi.org/10.1016/j.cedpsych.2019.101801>
- 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. <https://doi.org/10.1080/07370008.2013.769997>
- 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(4), 479–497. <https://doi.org/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(2), 171–185. <https://doi.org/10.1002/acp.3391>
- Mar, R. A., Li, J., Nguyen, A., & Ta, C. (2021). Memory and comprehension of narrative versus expository texts: A meta-analysis. *Psychonomic Bulletin & Review*, 28, 732–749. <https://doi.org/10.3758/s13423-020-01853-1>
- Mason, L., Scrimin, S., Tornatora, M. C., & Zaccoletti, S. (2017). Emotional reactivity and comprehension of multiple online texts. *Learning and Individual Differences*, 58, 10–21. <https://doi.org/10.1016/j.lindif.2017.07.002>
- Mazzocco, P. J., Green, M. C., Sasota, J. A., & Jones, N. W. (2010). This story is not for everyone: Transportability and narrative persuasion. *Social Psychological and Personality Science*, 1, 361–368. <https://doi.org/10.1177/1948550610376600>
- McCrudden, M. T. (2018). Text relevance and multiple source use. In J. L. G. Braasch, I. Bråten, & M. McCrudden (Eds.), *Handbook of multiple source use* (pp. 168–183). Routledge.

- McCrudden, M. T., & Barnes, A. (2016). Differences in student reasoning about belief-relevant arguments: A mixed methods study. *Metacognition and Learning*, 11(3), 275–303. <https://doi.org/10.1007/s11409-015-9148-0>
- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1–43. [https://doi.org/10.1207/s1532690xc1401\\_1](https://doi.org/10.1207/s1532690xc1401_1)
- McNamara, D. S., Ozuru, Y., & Floyd, R. G. (2011). Comprehension challenges in the fourth grade: The roles of text cohesion, text genre, and readers' prior knowledge. *International Electronic Journal of Elementary Education*, 4(1), 229–257.
- Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability and knowledge levels. *Cognition and Instruction*, 14(2), 139–179. [https://doi.org/10.1207/s1532690xc1402\\_1](https://doi.org/10.1207/s1532690xc1402_1)
- Murphy, P. K., & Alexander, P. A. (2004). Persuasion as a dynamic, multidimensional process: An investigation of individual and intra-individual differences. *American Educational Research Journal*, 41, 337–363. <https://doi.org/10.3102/00028312041002337>
- Nouwens, S., Groen, M. A., & Verhoeven, L. (2016). How storage and executive functions contribute to children's reading comprehension. *Learning and Individual Differences*, 47, 96–102. <https://doi.org/10.1016/j.lindif.2015.12.008>
- 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). Erlbaum.
- Pickren, S. E., Stacy, M., Del Tufo, S. N., Spencer, M., & Cutting, L. E. (2022). The contribution of text characteristics to reading comprehension: Investigating the influence of text emotionality. *Reading Research Quarterly*, 57(2), 649–667. <https://doi.org/10.1002/trq.431>
- RAND Reading Study Group, & Snow, C. (2002). *Reading for understanding: Toward an R&D program in Reading comprehension*. RAND Corporation. <http://www.jstor.org/stable/10.7249/mr14650eri>
- Rizzo, R. B. (2007). Individual differences in the analysis of informal reasoning fallacies. *Contemporary Educational Psychology*, 32, 459–484. <https://doi.org/10.1016/j.cedpsych.2007.01.001>
- 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). Information Age Publishing.
- Rouet, J.-F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, 52(3), 200–215. <https://doi.org/10.1080/00461520.2017.1329015>
- Rukavina, I., & Daneman, M. (1996). Integration and its effect on acquiring knowledge about competing scientific theories from text. *Journal of Educational Psychology*, 88, 272–287. <https://doi.org/10.1037/0022-0663.88.2.272>
- Sadoski, M., Goetz, E. T., & Rodriguez, M. (2000). Engaging texts: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology*, 92, 85–95. <https://doi.org/10.1037/0022-0663.92.1.85>
- Scharrer, L., Bromme, R., Britt, M. A., & Stadler, M. (2012). The seduction of easiness: How science depictions influence laypeople's reliance on their own evaluation of scientific information. *Learning and Instruction*, 22, 231–243. <https://doi.org/10.1016/j.learninstruc.2011.11.004>
- Schoor, C., Rouet, J.-F., Artelt, C., Mahlow, N., Hahnel, C., Kroehne, U., & Goldhammer, F. (2021). Readers' perceived task demands and their relation to multiple document comprehension strategies and outcome. *Learning and Individual Differences*, 88, 102018. <https://doi.org/10.1016/j.lindif.2021.102018>
- Schreiner, C., Appel, M., Isberner, M. B., & Richter, T. (2018). Argument strength and the persuasiveness of stories. *Discourse Processes*, 55(4), 371–386. <https://doi.org/10.1080/0163853X.2016.1257406>
- Shaw, F. W. (1996). The cognitive processes in informal reasoning. *Thinking and Reasoning*, 2, 51–80. <https://doi.org/10.1080/135467896394564>
- Siegrist, M., & Cvetkovich, G. (2001). Better negative than positive? Evidence of a bias for negative information about possible health dangers. *Risk Analysis*, 21(1), 199–206. <https://doi.org/10.1111/0272-4332.211102>
- Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition*, 20(5), 539–548. <https://doi.org/10.3758/BF03199586>
- Stadler, M., Scharrer, L., Skodzik, T., & Bromme, R. (2014). Comprehending multiple documents on scientific controversies: Effects of reading goals and signaling rhetorical relationships. *Discourse Processes*, 51(1–2), 93–116. <https://doi.org/10.1080/0163853X.2013.855535>
- Van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, 54(5–6), 360–372. <https://doi.org/10.1080/0163853X.2017.1306677>
- Van den Broek, P., Lorch, R. E. J., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory and Cognition*, 29(8), 1081–1087. <https://doi.org/10.3758/BF03206376>
- van Strien, J. L. H., Brand-Gruwel, S., & Boshuizen, H. P. A. (2014). Dealing with conflicting information from multiple nonlinear texts: Effects of prior attitudes. *Computers in Human Behavior*, 32, 101–111. <https://doi.org/10.1016/j.chb.2013.11.021>
- van Strien, J. L. H., Kammerer, Y., Brand-Gruwel, S., & Boshuizen, H. P. A. (2016). How attitude strength biases information processing and evaluation on the web. *Computers in Human Behavior*, 60, 245–252. <https://doi.org/10.1016/j.chb.2016.02.057>
- Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal*, 46, 1060–1106. <https://doi.org/10.3102/0002831209333183>
- Wiley, J., Jaeger, A. J., & Griffin, T. D. (2018). Effects of instructional conditions on comprehension from multiple sources in history and science. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 341–361). Routledge.
- Wiley, J., Steffens, B., Britt, M. A., & Griffin, T. D. (2014). Writing to learn from multiple-source inquiry activities in history. In G. Rijlaarsdam (Series Ed.), & P. Klein, P. Boscolo, C. Gelati, & L. Kilpatrick (Eds.), *Vol. 28, Writing as a learning activity* (pp. 120–148). Brill.
- Winkler, J. R., Appel, M., Schmidt, M. C. R., & Richter, T. (2022). The experience of emotional shifts in narrative persuasion. *Media Psychology*, 26(2), 141–171. <https://doi.org/10.1080/15213269.2022.2103711>
- Wolfe, M. B. W., & Mienko, J. A. (2007). Learning and memory of factual content from narrative and expository text. *British Journal of Educational Psychology*, 77(3), 541–564. <https://doi.org/10.1348/000709906X143902>
- Wolfe, M. B. W., & Williams, T. J. (2017). Effects of text content and beliefs on informal argument evaluation. *Discourse Processes*, 54, 1–17. <https://doi.org/10.1080/0163853X.2017.1319654>
- Wolfe, M. B. W., & Woodwyk, J. M. (2010). Processing and memory of information presented in narrative or expository texts. *British Journal of Educational Psychology*, 80, 341–362. <https://doi.org/10.1348/000709910x485700>

Submitted November 23, 2023  
Final revision received June 9, 2024  
Accepted July 15, 2024

**Sylvia M. Savvidou** (corresponding author) is a Special Scientist of Department of Psychology, University of Cyprus, Nicosia, Cyprus; email: [savvidousylvia@gmail.com](mailto:savvidousylvia@gmail.com).

**Irene-Anna Diakidoy** is a Professor of Department of Psychology, University of Cyprus, Nicosia, Cyprus; email: [diakidoy.irene-anna@ucy.ac.cy](mailto:diakidoy.irene-anna@ucy.ac.cy).

**Lucia Mason** is a Professor of Department of Developmental Psychology and Socialization, University of Padova, Padova, Italy; email: [lucia.mason@unipd.it](mailto:lucia.mason@unipd.it).