Memory & Cognition (2019) 47:33–46
https://doi.org/10.3758/s13421-018-0848-y

Knowledge revision through the lenses of the three-pronged approach

Panayiota Kendeou[1] · Reese Butterfuss[1] · Jasmine Kim[1] · Martin Van Boekel[1]

Published online: 16 August 2018

Abstract
In the present study, we employed the three-pronged approach to determine the *actual cognitive processes* theorized in knowledge revision. First, the Knowledge Revision Components (KReC) framework was identified as the guiding theory. Second, think-aloud analysis highlighted at which points in refutation texts readers detected discrepancies between their incorrect, commonsense beliefs and the correct beliefs, and the exact processes with which they dealt with these discrepancies—successfully or unsuccessfully, as indicated by posttest scores. Third, corroborating reading-time data and posttest data demonstrated that the structure of the refutation texts facilitated the coactivation and integration of the explanation with the commonsense belief, resulting in knowledge revision. Finally, an analysis directly connected the processes identified during think-aloud to sentence reading times. These findings systematically identify the cognitive processes theorized during knowledge revision and, in doing so, provide evidence for the conditions for revision outlined in the KReC framework.

Keywords Knowledge revision · Refutation texts · Three-pronged approach · Think-aloud · Cognitive processes

The efforts of researchers to identify ways to correct intuitions (Shtulman & Harrington, 2016), misconceptions (Fazio, Brashier, Payne, & Marsh, 2015; Potvin & Cyr, 2017), and misinformation (Ecker, Hogan, & Lewandowsky, 2017; Rapp, Hinze, Kohlhepp, & Ryskin, 2014) have intensified over the past few years. These efforts often involve presenting individuals with written accounts that challenge their intuitions or preexisting knowledge and beliefs. To date, many studies have used this approach and focused on identifying the factors that facilitate revision (Sinatra & Broughton, 2011; Tippett, 2010). This work, however, provided limited insights into readers' cognitive processing during revision (Lewandowsky & Oberauer, 2016). The Knowledge Revision Components framework (KReC; Kendeou & O'Brien, 2014) was proposed to fill this gap and provide one account for the revision processes during reading. As we will advocate next, while the findings from previous studies in the context of KReC are consistent with the processes *theorized* in the framework, they provide limited information about the *actual cognitive processes* at work during knowledge revision. In the present study, we address this issue by employing the three-pronged method (Magliano & Graesser, 1991). The three-pronged method is a rigorous and systematic study of discourse processes that requires linking theory to converging sources of empirical evidence in order to make claims about the processes being studied (Graesser, Swamer, & Hu, 1997; Magliano, Baggett, Johnson, & Graesser, 1993).

✉ Panayiota Kendeou
kend0040@umn.edu

[1] Department of Educational Psychology, University of Minnesota, 56 East River Road, Minneapolis, MN 55455, USA

The Knowledge Revision Components (KReC) framework

Several models and frameworks have been proposed to account for knowledge revision primarily in the conceptual change literature (e.g., Carey, 2000, 2009; Chi, 2008; Chinn & Brewer, 1993, 1998; Clark, 2006; diSessa, 2008; Dole & Sinatra, 1998; Hynd & Guzzetti, 1998; Ohlsson, 2009; Posner, Strike, Hewson, & Gertzog, 1982; Sinatra & Pintrich, 2003; Thagard, 2008; Vosniadou, 2008). The KReC framework (Kendeou & O'Brien, 2014) differs from these models in at least two ways. Specifically, knowledge revision, as articulated in KReC, differs from "change" in conceptual change models in that it does not assume an "erase and replace" mechanism. Instead, revision in the context of KReC is operationalized as the reduction of the reactivation of preexisting knowledge or misconceptions during reading. Second, KReC makes no a priori assumptions regarding the level at which knowledge is misconceived or the coherence of knowledge representations (Kendeou, Butterfuss, Van Boekel, & O'Brien, 2017). Rather, the emphasis is on identifying the conditions of knowledge revision during reading. The framework has been initially applied to the investigation of reading refutation texts, which are designed to promote knowledge revision by explicitly acknowledging incorrect beliefs about a topic, directly refuting them, and providing an accurate explanation (Hynd, 2001). Refutation texts have been found to support knowledge revision for elementary, secondary, and college-level students, both in the laboratory and classroom settings (Alvermann & Hague, 1989; Alvermann & Hynd, 1989; Braasch, Goldman, & Wiley, 2013; Guzzetti, Snyder, Glass, & Gamas, 1993; Hynd, Alvermann, & Qian, 1997; Lombardi, Danielson, & Young, 2016; Mason et al., 2017; Rapp & Kendeou, 2007, 2009; Sinatra & Broughton, 2011; Tippett, 2010; Van Loon, Dunlosky, Van Gog, Van Merriënboer, & De Bruin, 2015).

KReC identifies five principles to account for the knowledge revision process; the first two are *assumptions* (encoding and passive activation), and the remaining three are *conditions* necessary for the revision process (coactivation, integration, and competing activation). According to the encoding principle, once information has been encoded into long-term memory, it cannot be "erased" (e.g., Gillund & Shiffrin, 1984; Hintzman, 1986; Kintsch, 1988; Ratcliff, 1978; Ratcliff & McKoon, 1988). Thus, there is always the potential that it can be reactivated and influence comprehension. According to the passive activation principle, information in long-term memory is activated via passive memory processes (Myers & O'Brien, 1998; O'Brien & Myers, 1999). Because memory activation is both passive and unrestricted, any information that is related to the current contents of working memory has the potential to become activated, independent of whether it facilitates or interferes with comprehension. These two principles taken together raise an important question for knowledge revision: If incorrect knowledge cannot be erased and always has the potential to be reactivated and influence subsequent comprehension, then how can knowledge revision be accomplished?

The remaining three principles of KReC help answer this question by specifying the conditions that will reduce the reactivation, and thus the potential influence, of previously encoded incorrect knowledge. KReC operationalizes knowledge revision as the reduction of measurable activation of previously acquired misconceived knowledge. According to the coactivation principle, coactivation is a necessary condition for knowledge revision because it is the only way that new information can come in contact with previously encoded incorrect knowledge (Kendeou, Muis, & Fulton, 2011; Kendeou & van den Broek, 2007; van den Broek & Kendeou, 2008). According to the integration principle, knowledge revision can only occur when newly encoded information is integrated with this previously encoded misconceived knowledge. Any time new information is integrated with previously acquired information, the long-term memory representation of that information is updated to take into account the new information (O'Brien, Cook, & Gueraud, 2010; Zwaan & Madden, 2004). According to the competing activation principle, as the amount of newly encoded information is increased, it will begin to dominate the integrated network of information, draw increasing amounts of activation to itself, and at the same time, away from the previously acquired misconceived information. As activation is drawn away from misconceived information, the amount of interference from that information decreases accordingly, along with any measurable disruption in comprehension processes (e.g., Kendeou, Smith, & O'Brien, 2013; Kendeou, Walsh, Smith, & O'Brien, 2014; McNamara & McDaniel, 2004; Van Boekel, Lassonde, O'Brien, & Kendeou, 2017).

Initial evidence for the underlying processes theorized in KReC has been provided in a series of studies in the extant literature. In one study, Kendeou et al. (2014) tested the influence of the different components of refutation texts on knowledge revision. When participants read texts that contained a refutation-only, an explanation-only, or a refutation-plus-explanation elaboration of the correct information, the findings showed that all three textual elaborations (albeit to different degrees), drew more activation to the correct information and away from the misconception, thereby reducing the activation of the misconception during reading. However, only the refutation-plus-explanation condition resulted in detectable knowledge revision after reading the texts, and this effect was maintained even a month later, as detected by a delayed posttest. Kendeou et al. suggested that knowledge revision during reading is driven primarily by the competition between correct and incorrect information, with causal explanations increasing the probability that correct information "wins"' this competition. Importantly, in another series of experiments Kendeou et al. (2013) provided evidence that it is not merely the causal explanations per se that facilitate revision by winning the competition for reactivation but rather their high interconnectedness. Causal information inherently provides a rich set of connections to prior knowledge (Trabasso & Suh, 1993; Trabasso & van den Broek, 1985) and therefore an efficient and effective means of creating a network that will successfully compete for and draw activation, so that any interference from previously encoded misconceptions is reduced and/or eliminated.

Recently, the focus of this work has shifted to specific reader and text factors that can influence the knowledge revision process and specifically the competing activation of correct and incorrect information during reading. For example, Van Boekel et al. (2017) investigated the role of source credibility. Results revealed that source credibility influenced knowledge revision, but only when the source of the information was made salient via direct instructions to attend to sources prior to reading. In this case, high-credibility and low-credibility sources differentially influenced knowledge revision. Specifically, low-credibility sources disrupted knowledge revision processes and resulted in poorer performance on a posttest than high-credibility sources. In another study, Trevors, Kendeou, and Butterfuss (2017) investigated the moment-by-moment emotion processes during knowledge revision. Results showed that the emotion of *surprise* fluctuated in intensity at critical points during the knowledge revision process, suggesting the need to examine further the effects of emotions on competing activation and revision (see also Trevors, Muis, Pekrun, Sinatra, & Winne, 2016).

The three-pronged method

As noted at the beginning of this article, while the findings from previous work framed in KReC were consistent with the processes *theorized* in the context of the framework (coactivation, integration, and competing activation), they provided limited information about the *cognitive processes* at work during knowledge revision. Previous work, however, that examined the cognitive processes during reading of refutation texts is both relevant and informative. This work suggested that readers engage in cognitive conflict and monitoring when reading refutation texts (Ariasi & Mason, 2011; Broughton, Sinatra, & Reynolds, 2010; Kendeou et al., 2011; Kendeou & van den Broek, 2007; McCrudden & Kendeou, 2014), implicating possible coactivation of correct and incorrect knowledge during reading (van den Broek & Kendeou, 2008). These effects, however, were obtained with the use of one or two texts that often differed in structure across studies, limiting broad generalizations beyond a single topic or text. Most important, this work has informed the development of KReC rather than empirically test its theorized components and processes.

In the present study, we seek to systematically identify the processes theorized in KReC using the three-pronged method. The three-pronged method was proposed by Magliano and Graesser (1991) and has been used for decades to study inference generation in the field of discourse processes (e.g., Leon & Perez, 2001; Magliano et al., 1993; Magliano, Larson, Higgs, & Loschky, 2016; Sundermeier, van den Broek, & Zwaan, 2005; Trabasso & Magliano, 1996; Trabasso & Suh, 1993). The first prong requires researchers to use theory to guide experimentation and hypotheses generation. The second prong outlines the use of online measures of inference generation, such as verbal protocols, to test the predictions made in the first prong. The third prong stipulates the collection of corroborating online behavioral measures, such as reading times, to support the findings of the second prong (Graesser et al., 1997; Magliano et al., 1993; Magliano & Graesser, 1991). Thus, the three-pronged method ensures that the study of cognitive processing is both systematic and empirical (Magliano et al., 1993).

The present study

In line with the three-pronged method (Magliano & Graesser, 1991), the first prong of the present study involves using KReC to make predictions about the actual cognitive processes involved in knowledge revision during reading. The second prong involves the use of the think-aloud methodology in Experiment 1 to uncover the content of the cognitive processes in knowledge revision by observing participants' moment-by-moment processing. When only general prompts are provided (e.g., *What are you thinking right now?*) and readers have opportunities to practice and become familiar with the procedure, the use of think-alouds enables tracking of readers' cognitive processes without unduly influencing overall text comprehension (Afflerbach & Cho, 2009; Ericsson & Simon, 1993; Fletcher, 1986). Indeed, the methodology has been validated extensively in reading and discourse research (Afflerbach, 2002; Coté & Goldman, 1999; Magliano & Graesser, 1991; Magliano & Millis, 2003; Magliano, Trabasso, & Graesser, 1999; van den Broek, Kendeou, & White, 2008; Zwaan & Brown, 1996). The third prong involves the collection of corroborating behavioral evidence that supports the predictions and think-aloud findings. Therefore, in Experiment 2, we use reading times to provide further evidence of knowledge revision as conceptualized in KReC (i.e., reducing the reactivation of, and thus interference from, misconceptions). In both experiments, posttest data enable the examination of the situation models constructed after reading has been completed and provide converging evidence for revision.

It is important to note that even though KReC specifies processes that are not accessible to the reader (i.e., activation, integration, competing activation), the product of those processes (e.g., increased conflict as a result of coactivation and integration; reduced conflict as a result of competing activation) has the potential to be accessible and thus verbally reported during thinking aloud. If so, the contents of the think-aloud protocols should predict sentence reading times (Fletcher, 1986; Magliano & Graesser, 1991; Olson, Duffy, & Mack, 1984; Olson, Mack, & Duffy, 1981). To evaluate this claim, we also assess the relation between the processes identified during the think-aloud in Experiment 1 and the sentence reading times in Experiment 2, using the procedures suggested by Magliano and colleagues in the context of the three-pronged approach (Magliano & Graesser, 1991; Magliano et al., 1999). Further, to the extent that the think-aloud methodology itself does not alter naturalistic processes, the situation models constructed after reading in Experiments 1 and 2, and as a result readers' posttest scores, should not differ.

Experiment 1

The objective in Experiment 1 was to examine the extent to which the revision processes proposed by KReC are observed moment by moment during reading. Participants were asked to read a series of refutation and nonrefutation texts while performing a think-aloud task. After reading the set of texts, participants were asked to complete a comprehension posttest.

Each text included an introduction, an elaboration section (refutation plus explanation vs. control), a filler section, a correct outcome sentence, and a closing section (see example in the Appendix). Texts in the refutation elaboration condition included a *refutation section* that explicitly stated and refuted the target incorrect belief. The refutation was immediately followed by an *explanation section* that provided extra information supporting the correct belief. This was followed by a filler section that carried the story forward. After the filler section, *a correct outcome (target) sentence* was presented that stated the correct belief. Finally, all passages concluded with a closing section that wrapped up the storyline. Texts in the nonrefutation control condition simply continued the story line without refutation or explanation. In these texts, *a correct outcome (target) sentence* was also presented that stated the correct belief.

The first prong of the three-pronged approach requires that predictions are made using an a priori theory, in this case KReC. Several hypotheses can be drawn. We anticipated that when readers read the *refutation* that states and contradicts their incorrect prior knowledge, they would experience *cognitive conflict* as a result of the coactivation of correct and incorrect information, and likely engage in *comprehension monitoring* in an effort to establish coherence (Hypothesis 1). When readers proceed to read the *explanation,* we anticipated that the evoked *conflict and monitoring* from the preceding section would continue as readers attempt to establish coherence by encoding the interconnected explanation and *generating inferences* (Hypothesis 2). When readers read the *correct outcome sentence,* both the information associated with the incorrect belief and the explanation would compete for reactivation. Because the explanation is highly interconnected due to inherent connections to prior knowledge, it will "win" the competition for reactivation, drawing more activation to the correct information (and away from the incorrect belief) and thus facilitating the integration of the outcome sentence, reducing or eliminating *cognitive conflict*, while readers attempt to build coherence by generating more *inferences* (Hypothesis 3). In contrast, in the nonrefutation condition, reading of the *correct outcome sentence* will reactivate the incorrect belief because there is no causal explanation to compete for reactivation. As a result of the incorrect belief being reactivated, readers are expected to experience *cognitive conflict* and likely engage in *comprehension monitoring* in an effort to establish coherence (Hypothesis 4). No other concrete predictions can be made for reading the nonrefutation texts; assuming engagement with the reading task, readers will likely engage in typical comprehension processing as indicated by *inference* generation, as well as drawing *associations* to prior knowledge. Finally, to the extent that the aforementioned processes result in knowledge revision, posttest scores after reading should be significantly higher in the refutation text condition than scores in the nonrefutation text condition (Hypothesis 5).

Method

Participants

A total of 35 University of Minnesota undergraduate students enrolled in introductory psychology courses participated in the current study. Participants received partial course credit for their participation. Of the 35 participants, 20 were female and 15 were male, with an age range of 18 to 38 years (M = 20.29 years, SD = 3.50).

Design

There was one within-subjects factor, text type. Participants read five refutation (experimental) and five nonrefutation (control) texts. Cognitive processes were evaluated through the coding of the think-aloud protocols produced during reading (the coding scheme is detailed below). Knowledge revision was evaluated based on posttest accuracy scores.

Materials

Texts Texts consisted of 10 narrative-informational texts (Trevors et al., 2017), each addressing a common incorrect belief that was identified using a similar sample in previous studies (Van Boekel et al., 2017). Each text included an introduction, an elaboration section (refutation plus explanation vs. control), a filler section, a correct-outcome sentence, and a closing section (see the Appendix). All passages began with seven introductory sentences totaling 100 words, which served to establish the storyline. This was followed by one of two elaboration conditions: refutation plus explanation or control (nonrefutation plus nonexplanation). The refutation consisted of two sentences (33 words) that explicitly stated and refuted the target incorrect belief. The refutation was immediately followed by an explanation that provided extra information supporting the correct belief (average of 89.4

words, range: 84–95 words). The control section progressed the story, making no mention of either the incorrect belief or the refutation section (average of 88.3 words, range: 79–96 words). In both conditions, a filler section followed that carried the story forward. On average, the filler section consisted of three sentences and 49.7 words (range: 37–64 words). After the filler section, a correct outcome (target) sentence was presented that stated the correct belief. The correct outcome sentence averaged 18.9 words (range: 13–30 words). Finally, all passages concluded with a closing section that wrapped up the storyline (37 words). Each passage ended with a comprehension question that did not address information concerning the incorrect belief. Each text section was presented on its own card, resulting in eight cards per text.

Two material sets were constructed, each contained five texts in each of the two text type conditions. Across the two sets, each text occurred once in each of the two conditions. Participants were randomly assigned to a set.

Posttest The test included 10 two-tiered questions corresponding to the 10 incorrect beliefs targeted in the texts. The first tier was a typical true/false question, followed by the second tier requiring participants to provide an explanation for their response. Participants were awarded 1 point for each correct true/false question, and zero points for an incorrect answer. Correct explanations were awarded two points and incorrect explanations zero points. The reliability of the scores on the test was high ($\alpha = .84$).

Procedure

Participants were tested individually in a single session. The participants were informed that they were going to read some texts and think aloud after every section (Ericsson & Simon, 1993). The experimenter explained that thinking aloud is a process during which one is asked to state aloud his or her thoughts while reading a text. The participants were given a practice text and were asked to read each text section on the card aloud and talk about their thoughts. Reading aloud (instead of silently) was adopted because it allowed the researchers to identify where in the text the verbalizations were produced. There is precedent in the literature for using this approach (Bohn-Gettler & Kendeou, 2014; Trabasso & Suh, 1993; Kendeou & van den Broek, 2007; Magliano et al., 1999) for a practical reason: to accurately connect text to verbalizations. Participants were also instructed to make sure they understood what they were reading. If a participant had difficulty thinking aloud, the experimenter posed the following question: "What are you thinking right now?" When the practice text was completed, the participants were administered the entire set of texts, and they were asked to read and think aloud after every section. When the participants finished reading and thinking aloud, they were asked to complete the posttest.

Think-aloud coding

Participants' responses to each text section were transcribed. Each response was coded into a single category. These categories helped identify the type of process engaged in by a reader. The categories of responses were adapted from those used by Kendeou and van den Broek (2007), and included the following: *associations* (comments providing information activated from prior knowledge but not helpful for building coherence); *elaborative inferences—valid or invalid* (comments explaining the current text based on previous text and background knowledge); *text-based inferences* (comments accurately explaining the current text based on previous text); *paraphrases/summaries* (comments that capture the gist meaning of the current text); *cognitive conflict* (comments that express conflict between individuals' prior knowledge and the information in the current text); *monitoring comprehension* (comments that reflect individuals' understanding or lack thereof). Responses that did not fall into any of these categories were scored as *other*. Example coded responses are included in Table 1.

A second, independent coding was performed to identify the valence and activation dimensions of emotions. This coding pertained to a separate research aim, namely, the identification of moment-by-moment emotions during reading. More information about this coding and analyses can be found in Trevors et al. (2017).

All think-aloud protocols were scored by a trained undergraduate volunteer, and then a second trained rater (the last author) scored 11 of the think-alouds independently to verify consistency (*kappa* = 0.86).

Results and discussion

The analyses in both experiments were conducted using analysis of variance (ANOVA) with participants (F_1) and items (F_2) as random variables. For variables with skewness or kurtosis values > ± 1.5 (Gravetter & Wallnau, 2014), a related-samples Wilcoxon signed-rank test was conducted instead. In Experiment 1, we also conducted generalized linear mixed-models (GLMM) analyses with logit functions and in Experiment 2 linear mixed-models (LMM) analyses, with participants and items as random factors (Baayen, Davidson, & Bates, 2008; Richter, 2006). The lme4 package for R was used to conduct these analyses (Bates, Maechler, & Bolker, 2011). We present significance tests for the fixed effects of interest (z scores for GLMM analyses and t scores for LMM analyses). We analyzed the think-aloud responses after the refutation section, the explanation section, and the correct outcome sentence.

Table 1 Example think-aloud responses with coding for refutation and nonrefutation text conditions

Text section	Code and sample responses	
Refutation condition		
Refutation	*Cognitive conflict:* I always think a meteor would be hot, so I don't know what they meant by that.	*Paraphrasing:* Kate and Jerry are disagreeing. I am assuming there is going to be an argument.
Explanation	*Monitoring comprehension:* It is very interesting. I did not know that. I would probably not touch it.	*Text-based inference:* Even if I knew it was not hot, I would not decide to touch it either.
Correct outcome	*Elaborative inference—valid:* So, the meteors don't have any effect on the environment and won't damage [it].	*Paraphrasing:* Jerry is still explaining things. It is a little weird.
Nonrefutation condition		
Nonrefutation	*Association:* I do not know why he would have an astrophysics book because that sounds terrible to me.	*Monitoring comprehension:* That's interesting. I would try more to know about it [meteor] after seeing one.
Nonexplanation	*Association:* Well, Jerry sounds like my Uncle George, always offering facts and information.	*Elaborative inference—valid:* Surprising that it [meteor] did not hit anything. Just across the street from a house.
Correct outcome	*Cognitive conflict:* I thought that they were hot still from being in the atmosphere, but . . .	*Text-based inference:* So, Jerry definitely wants to find out the scientific reasons.

Refutation section

The results showed that there were significant differences between the refutation and the control texts with respect to associations, $F_1(1, 33) = 8.88, p = .005, \eta_p^2 = .21, F_2(1, 8) = 7.73, p = .024, \eta_p^2 = .49$ (GLMM analysis: $z = -2.72, p < .001$); text-based inferences, $F_1(1, 33) = 4.83, p = .035, \eta_p^2 = .13, F_2(1, 8) = 10.02, p = .013, \eta_p^2 = .55$ (GLMM analysis: $z = -2.17, p = .03$); elaborative inferences—valid, $F_1(1, 33) = 4.96, p = .033, \eta_p^2 = .13, F_2(1, 8) = 3.75, p = .089, \eta_p^2 = .31$ (GLMM analysis: $z = -2.26, p = .02$); cognitive conflict, Wilcoxon $z = -2.73, p = .006$ (GLMM analysis: $z = 3.22, p = .001$); and monitoring comprehension, $F_1(1, 33) = 16.38, p < .001, \eta_p^2 = .33, F_2(1, 8) = 14.87, p = .005, \eta_p^2 = .65$ (GLMM analysis: $z = 2.39, p = .02$). There were no differences with respect to elaborative inferences—invalid, Wilcoxon $z = -.44, p > .05$ (GLMM analysis: $z = .45, p > .05$) or paraphrases/summaries, Wilcoxon $z = -1.06, p > .05$ (GLMM analysis: $z = -.21, p > .05$).

As can be seen in Table 2, and consistent with Hypothesis 1, after reading the refutation sentences, participants reported significantly more cognitive conflict and monitoring in refutation texts than the control texts. In control texts, participants engaged in typical comprehension processing with minimal disruption (cognitive conflict), as indicated by significantly more text-based and valid elaborative inferences as well as associations.

Table 2 Think-aloud responses for the refutation versus nonrefutation (control) section in each condition

	Condition			
	Refutation		Control	
Processes	M	SD	M	SD
Associations*	.11	.18	.26	.28
Text-based inferences*	.15	.19	.26	.29
Elaborative inferences—valid*	.21	.24	.33	.26
Elaborative Inferences—invalid	.02	.06	.01	.05
Metacognitive Comments*	.25	.23	.08	.12
Cognitive Conflict*	.15	.18	.01	.03
Paraphrase/summary	.06	.17	.03	.11

*$p < .05$

Explanation section

The results showed that there were significant differences between the refutation and the control texts with respect to elaborative inferences—valid, $F_1(1, 33) = 7.90, p = .008, \eta_p^2 = .19, F_2(1, 8) = 7.01, p = .029, \eta_p^2 = .46$ (GLMM analysis: $z = -3.13, p = .002$); cognitive conflict,[1] Wilcoxon $z = -2.46, p = .014$; and monitoring comprehension, $F_1(1, 33) = 23.19, p < .001, \eta_p^2 = .41, F_2(1, 8) = 18.56, p = .003, \eta_p^2 = .69$ (GLMM analysis: $z = 4.62, p < .001$). There were no differences with respect to associations, $F_1(1, 33) = 3.55, p > .05, F_2(1, 8) = 3.93, p > .05$ (GLMM analysis: $z = -.79, p > .05$); text-based inferences, $F_1(1, 33) = 3.74, p > .05, F_2(1, 8) = 3.94, p > .05$ (GLMM analysis: z −1.90, $p > .05$); elaborative inferences—invalid, Wilcoxon $z = -.63$ $p > .05$ (GLMM analysis: $z = -.64, p > .05$); or paraphrases/summaries, Wilcoxon $z = -1.42, p > .05$ (GLMM analysis: $z = .36, p > .05$).

As can be seen in Table 3, after reading the explanation sentences, participants continued to report significantly more cognitive conflict and monitoring in refutation than the control texts, a finding consistent with Hypothesis 2. While reading the control texts, they also continued to engage in typical comprehension processing as indicated by significantly more valid elaborative inferences.

[1] GLMM analysis was not conducted because variability of cognitive conflict was limited in the control condition.

Table 3 Think-aloud responses for the explanation versus nonexplanation (control) section in each condition

	Condition			
	Refutation		Control	
Processes	M	SD	M	SD
Associations	.15	.18	.24	.28
Text-based inferences	.22	.22	.30	.27
Elaborative inferences—valid*	.13	.14	.25	.26
Elaborative inferences—invalid	.02	.06	.03	.08
Monitoring comprehension*	.35	.22	.11	.16
Cognitive conflict*	.05	.11	.00	.00
Paraphrase/summary	.07	.15	.04	.08

*$p < .05$

Correct outcome sentence

The results showed that there were significant differences between the refutation and the control texts with respect to text-based inferences, $F_1(1, 33) = 59.05$, $p < .001$, $\eta_p^2 = .64$, $F_2(1, 8) = 31.44$, $p = .001$, $\eta_p^2 = .79$ (GLMM analysis: $z = 5.80$, $p < .001$); elaborative inferences—valid, $F_1(1, 33) = 4.37$, $p = .044$, $\eta_p^2 = .12$, $F_2(1, 8) = 4.63$, $p = .064$, $\eta_p^2 = .37$ (GLMM analysis: $z = 2.70$, $p \leq .001$); monitoring comprehension, $F_1(1, 33) = 9.02$, $p = .005$, $\eta_p^2 = .22$, $F_2(1, 8) = 3.46$, $p = .100$, $\eta_p^2 = .30$ (GLMM analysis: $z = -2.70$, $p = .001$); and, cognitive conflict, Wilcoxon $z = -3.74$, $p < .001$ (GLMM analysis: $z = -1.73$, $p = .08$). There were no differences with respect to associations, Wilcoxon $z = -1.51$, $p > .05$ (GLMM analysis: $z = -1.34$, $p > .05$); elaborative inferences—invalid, Wilcoxon $z = -.27$, $p > .05$ (GLMM analysis: $z = -.90$, $p > .05$); or paraphrases/summaries, Wilcoxon $z = -.81$, $p > .05$ (GLMM analysis: $z = -.62$, $p > .05$).

As can be seen in Table 4, after reading the correct outcome sentence participants reported lower cognitive conflict and metacognitive comments in the refutation than control condition, suggesting that the competing activation (drawing more activation to the correct information and away from the incorrect belief) reduced or eliminated the reactivation of the incorrect belief in the refutation (consistent with Hypothesis 3), but not in the control condition (consistent with Hypothesis 4). Further, competing activation also resulted in increased coherence as reflected by higher frequency of coherence-building processes (i.e., inferences) in the refutation than the control condition.

Table 4 Think-aloud responses for the correct outcome sentence in each condition

	Condition			
	Refutation		Control	
Processes	M	SD	M	SD
Associations	.05	.10	.08	.14
Text-based inferences*	.47	.24	.18	.23
Elaborative inferences—valid*	.11	.11	.06	.10
Elaborative inferences—invalid	.04	.10	.05	.11
Monitoring comprehension*	.21	.20	.34	.25
Cognitive conflict*	.03	.09	.19	.17
Paraphrase/summary	.03	.09	.04	.11

*$p < .05$

Posttest

The results showed that items associated with reading refutation texts resulted in significantly higher scores ($M = 11.28$, $SD = 3.84$) than the control texts ($M = 7.54$, $SD = 3.09$), $F_1(1, 33) = 42.51$, $p < .001$, $\eta_p^2 = .56$, $F_2(1, 8) = 7.32$, $p = .027$, $\eta_p^2 = .48$, confirming the effectiveness of the refutation texts in facilitating knowledge revision (consistent with Hypothesis 5).

In summary, these results provided evidence for the specific hypotheses constructed in the context of the KReC framework regarding the processes of revision. Specifically, when texts contained refutations and explanations, participants initially reported higher levels of cognitive conflict and comprehension monitoring, likely the product of *coactivation and integration* of the correct and the incorrect beliefs. The mental representations readers built in this context competed with the incorrect beliefs for activation. To the extent that *competing activation* reduced the reactivation of the incorrect belief when reading the correct outcome sentence (as indicated by reduced conflict and increased valid inference making), the revision process was facilitated. The higher posttest scores after reading the refutation texts provided further evidence that this process was successful. In the control texts, despite attempts to respond to conflict (as indicated by cognitive conflict and monitoring comprehension), readers failed to engage in successful revision, as indicated by lower post-test scores.

Experiment 2

The objective in Experiment 2 was to obtain reading-time data that would corroborate the findings from Experiment 1, the third prong of the three-pronged method. In this experiment, a different sample of participants read the same refutation and nonrefutation texts used in Experiment 1, line by line, while their reading times were recorded. Specifically, we obtained reading times for the correct outcome sentence (see the Appendix for sample text). After completing the reading task, participants completed the same posttest used in Experiment 1.

Reading times are an unobtrusive and useful way of obtaining online evidence of knowledge revision (Kendeou et al., 2014). Following KReC, we anticipated that the reading times of the outcome sentence that stated the correct belief would be faster in the refutation condition than in the control condition, suggesting that the correct outcome sentence was integrated into the evolving mental text representation, while at the same time any interference from the incorrect belief was reduced or eliminated (Hypothesis 1). Also, to the extent that reading refutation texts resulted in knowledge revision and the think-aloud methodology in Experiment 1 did not alter naturalistic processing (and thus influenced situation model construction), posttest scores should be consistent with those in Experiment 1, namely significantly higher in the refutation text condition than in the control (Hypothesis 2).

Method

Participants

A total of 40 University of Minnesota undergraduate students enrolled in introductory psychology courses participated in the current study. Participants received partial course credit for their participation. Of the 40 participants, 24 were female and 16 were male, with an age range of 18 to 23 years (M = 19.23 years, SD = 1.23).

Design

There was one within-subjects factor, text type. Participants read five refutation (experimental) and five nonrefutation (control) texts. Two variables were obtained to measure knowledge revision: reading times of the correct outcome sentence and accuracy on the posttest.

Materials

The materials were the same as those in Experiment 1, with one modification. The texts were presented to participants line by line, in order to control for number of characters in the correct outcome sentences. In doing so, the filler sections were lengthened to set up a shorter correct outcome sentence. The filler sections now averaged four sentences and totaled 60 words. The correct outcome sentences were 39–41 characters long (M = 40.4). The reliability of the posttest scores on the test was high (α = .84).

Procedure

Participants read all of the texts on a computer screen. Participants were instructed to rest their preferred hand on the line-advance key (space bar). Each trial began with the word "READY" in the center of the screen. When participants were ready to read a passage, they pressed the line-advance key. Each press of the key erased the current line of text (approximately seven words) and presented the next line of text. Reading time was measured as the time between key presses. Participants were instructed to read at a normal and comfortable reading rate. Following the last line of each passage, participants were presented with a comprehension question to which participants responded by either pressing the "yes" or "no" key. On the trials in which participants made an error, the word "ERROR" appeared in the middle of the screen for 750 milliseconds. Before beginning the experimental passages, participants read two practice passages to ensure that they were familiar with and understood the procedure.

Upon completion of the reading task, participants completed the 10-item posttest. Lastly, participants completed a short demographic form and were debriefed and thanked for their participation in the study.

Results and discussion

To the extent that reading refutation texts influence knowledge revision, reading times of the correct outcome sentence in the refutation text condition should be faster than the nonrefutation control text condition. Faster reading times would indicate that the correct outcome sentence was integrated into the evolving mental representation while at the same time suggesting that any interference from the incorrect belief was reduced (Hypothesis 1). The findings provided support for this hypothesis and showed that the correct outcome sentences were read significantly faster in the refutation (M = 1751, SD = 476) than in the nonrefutation control condition (M = 2065, SD = 634), $F_1(2, 38) = 29.53$, $p < .001$, $\eta_p^2 = 0.44$; $F_2(1, 8) = 21.05$, $p = .002$, $\eta_p^2 = 0.72$ (LMM analysis: $t = -5.14$, $p < .05$).

The analysis of posttest scores showed that items associated with reading refutation texts resulted in significantly higher scores (M = 11.43, SD = 3.60) than items associated with the nonrefutation control texts (M = 5.60, SD = 3.23), $F_1(2, 42) = 678.61$, $p < .001$, $\eta_p^2 = 0.69$, $F_2(1, 8) = 222.04$, $p < .001$, $\eta_p^2 = 0.96$, confirming Hypothesis 2.

The reading time and posttest findings taken together indicate that reading refutation texts facilitated knowledge revision. These results demonstrate that the refutation and explanation sections are integrated into the evolving mental representation of the text, helping the newly acquired information to dominate when competing for activation with the incorrect belief. As a result, any interference from the incorrect belief is reduced, as reflected by faster reading times on the outcome sentence. These findings provide corroborating evidence for the findings obtained in Experiment 1.

Connecting Experiments 1 and 2

An advantage of using the three-pronged method is the opportunity to establish convergence among think-aloud processes and reading times. In this context, we look for evidence of convergence, direct and indirect, among the different data sources. Recall that in the think-aloud experiment (Experiment 1) participants' verbal protocols were analyzed at three important points, namely after reading the refutation, explanation, and target sentences. In the reading time experiment (Experiment 2), participants' reading times were recorded on the target sentences. Thus, the target sentences present a unique opportunity to directly examine the convergence between the think-aloud and reading-time data. Furthermore, in both experiments, participants completed the same posttest after either thinking aloud (Experiment 1) or reading silently (Experiment 2). Thus, posttest scores present another opportunity to directly examine convergence between the two experiments.

Think-alouds and reading times

As noted above, even though KReC specifies processes that are not accessible to the reader, the product of those processes has the potential to be accessible and thus verbally reported during thinking aloud. If so, the contents of the think-aloud protocols should predict sentence reading times (Fletcher, 1986; Magliano & Graesser, 1991). To evaluate this assumption, we followed the procedures suggested by Magliano and colleagues (Magliano et al., 1999) and conducted a multiple-regression analysis. Criterion variables consisted of the interaction between condition (refutation vs. control) and each of the cognitive processes that were produced after reading the target sentences in Experiment 1 (e.g., Condition × Comprehension Monitoring), whereas the dependent variable consisted of the average reading time for the target sentences recorded in Experiment 2. The model accounted for significant variance in sentence reading times, $F(4, 65) = 2.66$, $p = .040$; $R^2 = .14$. One interaction was responsible for this effect, Condition × Comprehension Monitoring. Sentence reading times increased as a function of the frequency of metacognitive comments produced in the control condition, $\beta = .385$, $t(65) = 2.61$, $p = .011$ (LMM, $t = -3.22$, $p = .001$).

This finding is consistent with the results of Experiment 1, demonstrating that metacognitive comments were higher in frequency in the control than the refutation condition after reading the target sentence (see Table 4), and with the results of Experiment 2 demonstrating that reading times of the target sentence were slower in the control than in the refutation condition. It follows that the reading time slowdown (in Experiment 2) is likely the product of participants' response to interference as reflected by metacognitive comments (in Experiment 1).

In line with the recommendations made by Magliano and Graesser (1991), the convergence of the think-aloud and reading-time data at the target sentence provides direct evidence for the processes theorized in KReC and inferred via the think-alouds. This convergence also provides indirect evidence for the processes inferred at the refutation and explanation sentences. Because reading times were not recorded on those sentences, obtaining direct evidence was not possible.

Posttest scores

The processes that unfold during reading influence directly the construction of the situation model after reading has been completed (Kendeou & O'Brien, 2018; Rapp & Mensink, 2011; van den Broek & Helder, 2017). If the think-aloud methodology changed naturalistic processing, then situation model construction, and thus posttest scores in Experiments 1 and 2 would differ. Alternatively, if the influence of the think-aloud methodology was minimal, then situation model construction, and thus post-test scores in Experiments 1 and 2 would not differ. The results of an ANOVA on posttest scores with experiment as a between-subjects factor and text condition as a within-subjects factor showed that scores in Experiment 1 ($M = 8.51$, $SE = .45$) were not significantly different than scores in Experiment 2 ($M = 9.41$, $SE = .48$), $F(1, 73) = 1.82$, $p > .05$. As expected, scores with items associated with the refutation condition ($M = 11.35$, $SE = .43$) were higher than the control condition, ($M = 6.57$, $SE = .36$), $F(1, 73) = 119.20$, $p < .001$, $\eta_p^2 = 0.62$, across both experiments.

In line with the recommendations made by Magliano and Graesser (1991), the convergence of the posttest scores across the two experiments provides direct evidence that the think-aloud methodology had a rather minimal impact on readers' naturalistic processes during reading. This convergence also provides indirect evidence that the processes theorized in KReC and inferred via the think-alouds were the result of revision and not an artifact of the think-aloud methodology.

General discussion

In the present study, we employed the three-pronged approach (Magliano & Graesser, 1991), a rigorous and systematic method for studying discourse processes, to provide evidence for the *cognitive processes* theorized in KReC (Kendeou & O'Brien, 2014). KReC proposes that knowledge revision can be facilitated under three conditions: (a) coactivation of correct and incorrect information, subsequent (b) integration into a single mental network with a supporting, highly interconnected explanation of the correct information, which in turn can (c) compete and "win" the activation while simultaneously reducing the activation of and interference from incorrect information. While the findings from previous studies in the context of

KReC were consistent with the processes theorized in the framework, they provided limited information about the cognitive processes at work during knowledge revision.

Using a think-aloud methodology in Experiment 1, we provided evidence that after reading the refutation and explanation sections of the refutation texts, readers' verbal reports reflected processes of cognitive conflict and comprehension monitoring. These processes are consistent with the products of two of the conditions for revision outlined in KReC, *coactivation* of correct and incorrect information and *integration* of that information into the text mental representation. Relatedly, after reading the correct outcome sentence in the refutation texts, readers reported decreased conflict and comprehension monitoring and increased engagement in inference making (when compared with nonrefutation control texts). This finding is consistent with the product of the third condition of KReC, *competing activation*, suggesting that the highly interconnected causal network of the explanation in the refutation texts dominates the representation and reduces disruption caused by the incorrect information. This network consists of text-based and knowledge-based inferences readers encode during reading, as evidenced by the increased frequency of these processes in readers' verbal reports. Taken together, the orchestration of these processes results in detectable knowledge revision after reading has been completed, as reflected by posttest scores, showing close association between the online processes and offline products of reading comprehension (Kendeou & O'Brien, 2018; Rapp & Mensink, 2011; van den Broek & Helder, 2017). Specifically, readers' posttest scores in Experiment 1 for items associated with refutation texts were significantly higher than those associated with the nonrefutation control texts.

Using a reading time methodology in Experiment 2, we provided converging evidence for the KReC's main revision mechanism, competing activation. Specifically, the highly interconnected explanation in the refutation condition "wins" the competition for reactivation during reading of the correct outcome sentence, as indicated by faster reading times in the refutation than in the control condition. These faster reading times are taken as an indication of easier integration of the correct outcome sentence in the evolving mental representation of the text. At the same time, faster reading times also suggest no measurable disruption, and thus reduced reactivation of the incorrect information. In the context of KReC, reducing the reactivation of the incorrect belief is measurable evidence for knowledge revision (Kendeou et al., 2014). Converging evidence for successful knowledge revision was obtained from posttest scores. Specifically, readers' posttest scores in Experiment 2 for items associated with refutation texts were significantly higher than those associated with the nonrefutation control texts.

The processes of coactivation, integration, and competing activation theorized in KReC are not accessible to the reader. However, the product of those processes ought to be accessible. In the present study, we directly connected processes inferred during a think-aloud and a sentence reading time methodology (Magliano & Graesser, 1991). The results from this analysis demonstrated that the reading time slow-down on the correct outcome sentence obtained in the nonrefutation control condition in Experiment 2 was directly predicted by increased frequency of metacognitive comments in Experiment 1. These metacognitive comments may also be the product of validation processes. Validation is often defined as the monitoring of information consistency during comprehension (e.g., Richter, 2015; Singer, 2013). Readers routinely validate information against prior knowledge or prior text without being instructed to do so (Isberner & Richter, 2014; Richter & Maier, 2017). Presumably, validation follows the activation and integration of information during comprehension in a parallel, asynchronous fashion (Cook & O'Brien, 2014). Thus, validation has the potential to contribute to knowledge revision or its failure (Kendeou, 2014). However, methodologies with higher spatial and temporal precision (e.g., eye tracking) than those used in the present study will be needed to examine this issue.

The present study also demonstrates how the use of the three-pronged method (Magliano & Graesser, 1991) can be applied in the context of knowledge revision. Combining think-aloud and sentence-by-sentence reading methodologies is important because each of these has strengths and limitations. Think-aloud procedures provide rich information regarding the specific content of processing, but they always run the risk of changing readers' naturalistic processing and overall comprehension (Fox, Ericsson, & Best, 2011; Schooler, 2011). To reduce the probability of changing readers' comprehension, we provided general prompts (e.g., *What are you thinking right now?*) and opportunities to practice thinking aloud before reading the texts. Thus, when evaluating the processes inferred through think-aloud, one needs converging evidence to ensure that the self-reported reading activities are not elicited by the procedure itself and hence are part of spontaneous reading (for a review, see Ericsson & Simon, 1993). This converging evidence can be obtained via a less obtrusive method, such as sentence reading times. Sentence-by-sentence reading is sensitive to initial stages of comprehension and memory access. Because each sentence is presented in isolation, any disruption in comprehension from earlier portions of the text can only be resolved through memory processes. Combining these methodologies with careful text design allows for both the description of the actual processes during reading and the evaluation of the extent to which they occur moment-by-moment. We demonstrated the convergence of these processes in the present investigation by connecting think-aloud and reading-time data obtained from two independent samples. Indirect evidence was obtained by qualitatively comparing the patterns of results across the two experiments, whereas direct evidence was obtained by predicting reading times from think-aloud data on the correct outcome sentence and statistically comparing posttest scores across experiments. Direct and indirect evidence suggests convergence and rather minimal influences of the think-aloud methodology on the cognitive processes under investigation.

One limitation of the current study is the absence of a pretest measure of prior knowledge. This may raise concerns as to whether participants held the targeted commonsense, incorrect beliefs or misconceptions. Even though we cannot be certain that participants possess *all* of the misconceptions targeted in the present study, evidence from the think-aloud experiment suggests that at least, on average, these misconceptions are prevalent in the targeted population. We observed that participants engaged in cognitive conflict and comprehension monitoring when coactivation of correct and incorrect beliefs was hypothesized to be taking place (after the refutation in the refutation text and after the correct outcome sentence in the nonrefutation text). If participants did not hold the incorrect beliefs, there would be no need to engage in either of these processes. Furthermore, that readers who received nonrefutation texts endorsed the incorrect beliefs at higher rates (lower posttest scores) provides converging evidence that readers held the misconceptions initially, and revision was not facilitated in that condition. Perhaps more convincing, the results from a pretest in a recent study that used the same texts and sampled from the same population as the present study provided direct evidence for the high prevalence (range: 72%–95%) of these misconceptions (Van Boekel et al., 2017).

A second limitation concerns the use of an immediate posttest as a measure of knowledge revision. Specifically, it could be argued that participants' performance was a function of merely answering questions based on the texts they have just read, and not because they have engaged in knowledge revision per se. Previous work, however, has shown that in some cases the benefits of reading refutation texts were maintained even a month after reading when a delayed posttest was also used (Kendeou et al., 2014). The delayed effects, though, may have been driven by a "testing effect," namely, subsequent learning of information that results from having been tested on that information (McDaniel, Anderson, Derbish, & Morrisette, 2007). To address this issue, Walsh, Kendeou, and O'Brien (2015) systematically varied the time and frequency of testing after reading refutation texts, such that students received a delayed posttest with or without immediate posttest. Results showed that the testing effect did not play a significant role in posttest performance, thereby indicating that the long-term benefit in performance was the result of reading refutation texts. It is important to note, though, that long-term benefits from reading refutation texts are rather rare and depend highly on the type of misconceptions being targeted.

In conclusion, the present set of experiments adds to our understanding of the actual cognitive processes that facilitate knowledge revision during reading. The results demonstrate that the conditions for knowledge revision theorized in the KReC framework result in measurable processing differences during reading of refutation and nonrefutation texts, respectively. From a theory construction perspective, these results can help add to what we know about how readers respond when they are confronted with discrepant information, and the conditions under which these responses may facilitate knowledge revision as theorized in KReC. Specifically, the findings demonstrate that when textual information is strengthened in memory, it has the potential to dominate and eliminate measurable disruption caused by incorrect information in prior knowledge (i.e., misconceptions). Open questions aside, our results indicate that in the absence of such textual information, misconceptions persist and continue to influence both reasoning and performance.

Acknowledgements The authors would like to thank Jean-Baptiste Quillien and Isabella Albright for their assistance with data collection and coding. Preparation of this manuscript was supported by the Guy Bond Endowed Chair in Reading.

Appendix

Example text[2]

Introduction

After a busy day at work, Kate was out for her nightly run. About halfway through the run, she stopped at a corner to rest and stretch. Kate looked up at the clear night sky while she took a sip from her water bottle. She saw a meteor falling beyond the trees and watched until it hit the ground. She quickly ran about 400 yards to the site where the meteor landed. When she arrived there were already several people there. She noticed that her neighbor Jerry had also come down the street to see what was going on.

Refutation section

Kate warned everyone not to touch the meteor because it would be hot and they could get burned. However, Jerry said that they should not worry because it actually should not be hot.

[2] Refutation condition includes refutation and explanation sections; non-refutation control condition includes non-refutation and non-explanation sections.

Nonrefutation (control) section

Kate was excited and curious because she had never seen a meteor on the ground before. Jerry said that he could look up more about meteors in the astrophysics book that he had.

Explanation section

He explained that the high speed of the meteor when it enters the atmosphere causes it to melt or vaporize its outermost layer. The hot molten layer quickly blows off and the inside of the meteor does not have time to heat up again before passing through the atmosphere. This is because meteors are poor conductors of heat. Jerry told the crowd that many meteors that make it to Earth are actually found covered in frost. Despite this information, they all decided it was still a good idea not to touch it.

Nonexplanation (control) section

He told them that he had always been very interested in space and had read many articles about the research that they had been conducting in the space program. Jerry was known for offering up facts and information to anyone that would listen to him. He was sure that his book will have all sorts of facts about meteors. He walked across the street to get the book from his house while more people gathered around the meteor. They could not believe a meteor landed in their very own town.

Filler

Kate continued to stare at the meteor. She had never seen anything like this in person before, and figured that would be true of many people here. What if a television crew came to interview witnesses? She could be on TV! She had to come across as smart if she was interviewed.

Correct outcome sentence

She listened carefully as Jerry assured everyone that *meteors landing on Earth are always cold*.

Closing section

Police cars were now starting to arrive. The police told the crowd they had to go home because they needed to block off the area. Kate decided to sprint home to tell her family about the news.

References

Afflerbach, P. (2002). Verbal reports and protocol analysis. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), Handbook of reading research (Vol. 3, pp. 87–103). Mahwah: Erlbaum.

Afflerbach, P., & Cho, B.-Y. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. E. Israel & G. G. Duffy (Eds.), Handbook of research on reading comprehension (pp. 69–90). New York: Routledge.

Alvermann, D. E., & Hague, S. A. (1989). Comprehension of counterintuitive science text: Effects of prior knowledge and text structure. *The Journal of Educational Research, 82*(4), 197–202.

Alvermann, D. E., & Hynd, C. R. (1989). Effects of prior knowledge activation modes and text structure on nonscience majors' comprehension of physics. *The Journal of Educational Research, 83*(2), 97–102.

Ariasi, N., & Mason, L. (2011). Uncovering the effect of text structure in learning from a science text: An eye-tracking study. *Instructional Science*, *39*(5), 581–601.

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

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes [Computer software]. Vienna: R Foundation for Statistical Computing. Retrieved from http://CRAN.R-project.org/package-lme4

Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology, 39,* 206–219.

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

Broughton, S. H., Sinatra, G. M., & Reynolds, R. E. (2010). The nature of the refutation text effect: An investigation of attention allocation. *The Journal of Educational Research*, *103*(6), 407–423.

Carey, S. (2000). Science education as conceptual change. *Journal of Applied Developmental Psychology*, *21*(1), 13–19.

Carey, S. (2009). The origin of concepts. Oxford: Oxford University Press.

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), International handbook of research on conceptual change (pp. 61–82). New York: Taylor & Francis.

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

Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching, 35*(6), 623–654.

Clark, D. B. (2006). Longitudinal conceptual change in students' understanding of thermal equilibrium: An examination of the process of conceptual restructuring. *Cognition and Instruction, 24*(4), 467–563.

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes*, *51*(1/2), 26–49.

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: Erlbaum.

diSessa, A. A. (2008). Can students re-invent fundamental scientific principles? Evaluating the promise of new-media literacies. In T.

Willoughby & E. Wood (Eds.), Children's learning in a digital world (pp. 218–248). Malden: Blackwell Publishing.

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist, 33*(2/3), 109–128.

Ecker, U. K., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and repetition of misinformation: Helping or hindering its retraction? *Journal of Applied Research in Memory and Cognition, 6*(3), 185–192.

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

Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General, 144*(5), 993–1002.

Fletcher, C. R. (1986). Strategies for the allocation of short-term memory during comprehension. *Journal of Memory & Language*, *25*, 43–58.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin, 137*, 316–344.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*(1), 1–67.

Graesser, A. C., Swamer, S. S., & Hu, X. (1997). Quantitative discourse psychology. *Discourse Processes*, *23*(3), 229–263.

Gravetter, F., & Wallnau, L. (2014). Essentials of statistics for the behavioral sciences (8th). Belmont: Wadsworth.

Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly, 28*(2), 117–159.

Hintzman, D. L. (1986). 'Schema abstraction' in a multiple-trace memory model. *Psychological Review, 93*(4), 411–428.

Hynd, C. (2001). Refutational texts and the change process. *International Journal of Educational Research, 35*(7/8), 699–714.

Hynd, C., Alvermann, D., & Qian, G. (1997). Preservice elementary school teachers' conceptual change about projectile motion: Refutation text, demonstration, affective factors, and relevance. *Science Education, 81*(1), 1–27.

Hynd, C., & Guzzetti, B. J. (1998). When knowledge contradicts intuition: Conceptual change. In C. Hynd (Ed.), Learning from text across conceptual domains (pp. 139–164). Mahwah: Erlbaum.

Isberner, M.-B., & Richter, T. (2014). Does validation during language comprehension depend on an evaluative mindset? *Discourse Processes*, *51*(1/2), 7–25.

Kendeou, P. (2014). Validation and comprehension: An integrated overview. *Discourse Processes*, *51*(1/2), 189–200.

Kendeou, P., & Butterfuss, R., Van Boekel, M., & O'Brien, E. J. (2017). Integrating relational reasoning and knowledge revision during reading. *Educational Psychology Review,* 1-13

Kendeou, P., Muis, K., & Fulton, S. (2011). Reader and text factors in reading comprehension. *Journal of Research in Reading, 34*, 365–383.

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

Kendeou, P., & O'Brien, E. J. (2018). Theories of text processing: A view from the top-down. In M. Schober, D. N. Rapp, & M. A. Britt (Eds.), Handbook of discourse processes (2nd, pp. 7–21). New York: Routledge.

Kendeou, P., Smith, E., & O'Brien, E. (2013). Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 854–865.

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.

Kendeou, P., Walsh, E., Smith, E., & O'Brien, E. (2014). Knowledge revision processes in refutation texts. *Discourse Processes, 51*(5/6), 374–397.

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

Leon, J., & Perez, O. (2001). The influence of prior knowledge on time course of clinical diagnosis inferences: A comparison of experts and novices. *Discourse Processes, 31*(2), 187–213.

Lewandowsky, S., & Oberauer, K. (2016). Motivated rejection of science. *Current Directions in Psychological Science, 25*, 217–222.

Lombardi, D., Danielson, R. W., & Young, N. (2016). A plausible connection: Models examining the relations between evaluation, plausibility, and the refutation text effect. *Learning & Instruction, 44,* 74–86.

Magliano, J., Larson, A., Higgs, K., & Loschky, L. (2016). The role of linguistic and grounded systems in generating inferences while processing visual narratives. *Memory & Cognition, 44*(2), 207–219.

Magliano, J. P., Baggett, W. B., Johnson, B. K., & Graesser, A. C. (1993). The time course of generating causal antecedent and causal consequence inferences. *Discourse Processes*, *16*(1/2), 35–53.

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics*, *20*(3), 193–232.

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

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

Mason, L., Baldi, R., Di Ronco, S., Scrimin, S., Danielson, R. W., & Sinatra, G. M. (2017). Textual and graphical refutations: Effects on conceptual change learning. *Contemporary Educational Psychology, 49,* 275–288.

McCrudden, M. T., & Kendeou, P. (2014). Exploring the link between cognitive processes and learning from refutational text. *Journal of Research in Reading, 37*, 116–140.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19*(4/5), 494–513.

McNamara, D. S., & McDaniel, M. (2004). Suppressing irrelevant information: Knowledge activation or inhibition? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30,* 465–482.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*(2/3), 131–157.

O'Brien, E. J., Cook, A. E., & Gueraud, S. (2010). Accessibility of outdated information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 979–991.

O'Brien, E. J., & Myers, J. L. (1999). Text comprehension: A view from the bottom up. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso (pp. 35–53). Mahwah: Erlbaum.

Ohlsson, S. (2009). Resubsumption: A possible mechanism for conceptual change and belief revision. *Educational Psychologist, 44*, 20–40.

Olson, G., Duffy, S., & Mack, R. (1984). Thinking-out-loud as a method for studying real time comprehension processes. In D. E. Kieras (Ed.), New methods in reading comprehension research (pp. 253–286). Hillsdale: Erlbaum.

Olson, G. M., Mack, R. L., & Duffy, S. A. (1981). Cognitive aspects of genre. *Poetics*, *10*(2/3), 283–315.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211–227.

Potvin, P. & Cyr, G. (2017). Toward a durable prevalence of scientific conceptions: Tracking the effects of two interfering misconceptions about buoyancy from preschoolers to science teachers. *Journal of Research in Science Teaching, 54*(9), 1121–1142. doi:https://doi.org/10.1002/tea.21396

Rapp, D. N., Hinze, S. R., Kohlhepp, K., & Ryskin, R. A. (2014). Reducing reliance on inaccurate information. *Memory & Cognition, 42*, 11–26.

Rapp, D. N., & Kendeou, P. (2007). Revising what readers know: Updating text representations during narrative comprehension. *Memory & Cognition*, *35*(8), 2019–2032.

Rapp, D. N., & Kendeou, P. (2009). Noticing and revising discrepancies as texts unfold. *Discourse Processes*, *46*(1), 1–24.

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

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95*, 385–408.

Richter, T. (2006). What is wrong with ANOVA and multiple regression? Analyzing sentence reading times with hierarchical linear models. *Discourse Processes, 41,* 221–250.

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes, 52*, 337–352.

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist, 52*, 148–166.

Schooler, J. W. (2011). Introspecting in the Spirit of William James: Comment on Fox, Ericsson, and Best (2011). *Psychological Bulletin, 137*(2), 345–350.

Shtulman, A., & Harrington, K. (2016). Tensions between science and intuition across the lifespan. *Topics in Cognitive Science, 8*(1), 118–137.

Sinatra, G., & Broughton, S. (2011). Bridging reading comprehension and conceptual change in science education: The promise of refutation text. *Reading Research Quarterly, 46*(4), 374–393.

Sinatra, G. M., & Pintrich, P. R. (Eds.). (2003). Intentional conceptual change. Mahwah: Erlbaum.

Singer, M. (2013). Validation in reading comprehension. *Current Directions in Psychological Science*, *22*(5), 361–366.

Sundermeier, B. A., van den Broek, P., & Zwaan, R. A. (2005). Causal coherence and the availability of locations and objects during narrative comprehension. *Memory & Cognition*, *33*(3), 462–470.

Thagard, P. (2008). Explanatory coherence. In J. E. Adler & L. J. Rips (Eds.), Reasoning: Studies of human inference and its foundations (pp. 471–513). New York: Cambridge University Press.

Tippett, C. D. (2010). Refutation text in science education: A review of two decades of research. *International Journal of Science and Mathematics Education, 8*(6), 951–970.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes*, *21*(3), 255–287.

Trabasso, T., & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes*, *16*(1/2), 3–34.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24*(5), 612–630.

Trevors, G. J., Kendeou, P., & Butterfuss, R. (2017). Emotion processes in knowledge revision. *Discourse Processes*, *54*(5/6), 406–426. doi: https://doi.org/10.1080/0163853X.2017.1312201

Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Winne, P. H. (2016). Identity and epistemic emotions during knowledge revision: A potential account for the backfire effect. *Discourse Processes*, *53*, 339–370.

Van Boekel, M., Lassonde, K., O'Brien, E. J., & Kendeou, P. (2017). Source credibility and the processing of refutation texts. *Memory & Cognition, 45,* 168–181.

van den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology, 22*(3), 335-351.

van den Broek, P., Kendeou, P., & White, M. J. (2008). Cognitive processes during reading: Implications for the use of multimedia to foster reading comprehension. In A. G. Bus & S. B. Neuman (Eds.), Multimedia and literacy development: Improving achievement for young learners (pp. 57–73). New York: Rutledge.

van den Broek, P. W., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes 54*(5/6), 360–372.

Van Loon, M. H., Dunlosky, J., Van Gog, T., Van Merriënboer, J. J. G., & De Bruin, A. B. H. (2015). Refutations in science texts lead to hypercorrection of misconceptions held with high confidence. *Contemporary Educational Psychology, 42,* 39–48.

Vosniadou, S. (2008). International handbook of research on conceptual change. London: Routledge.

Walsh, E., Kendeou, P., & O'Brien, E. J. (2015, July). The long-term benefit of refutation text on knowledge revision: Not just a testing effect. Poster presented at the Annual Meeting of the Society for Text and Discourse, Minneapolis.

Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, *21*, 289–327.

Zwaan, R. A., & Madden, C. (2004). Updating situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 283–288.