The effects of refutation texts on generating explanations

Kelsey K. Willb,⁎, Andalusia Masada, Haley A. Vlachb, Panayiota Kendeoua

a University of Minnesota, Department of Educational Psychology, 56 E River Road, Minneapolis, MN 55455, United States of America
b University of Wisconsin-Madison, Department of Educational Psychology, 859 Education Sciences Building, 1025 W. Johnson Street, Madison, WI 53706, United States of America

ARTICLE INFO

Keywords:
Explanation
Misconception
Comprehension
Knowledge revision

ABSTRACT

The aim of the present study was to examine the effect of refutation texts on generating explanations and explore individual differences in quality of explanations. Seventy-five undergraduates read refutation and non-refutation texts addressing common scientific misconceptions and completed a two-tier post-test including True/False and open-ended questions. Readers’ written post-test explanations were coded on accuracy and quality dimensions. The analysis of these explanations showed less circularity and uncertainty descriptors and more accuracy and causal connections in the refutation than the non-refutation condition. Further, three distinct clusters of readers emerged: coherence-building readers, non-coherence building readers, and promiscuous readers. Results demonstrate examining explanation characteristics can provide a useful tool for revealing how readers’ knowledge influences learning from texts.

1. Introduction

When students enter learning environments they may bring along incorrect knowledge about different topics, making it difficult to acquire new related knowledge (e.g., Carey, 2009; Chi, 2005; Clement, 1991; Kendeou, Walsh, Smith, & O’Brien, 2014; Novak, 1988; Posner, Strike, Hewson, & Gertzog, 1982; Trumper, 2001; Vosniadou & Brewer, 1992, 1994). Misconceptions of scientific knowledge occur when one has inaccurate knowledge about phenomena and concepts (Kendeou & van den Broek, 2005). Decades of research have demonstrated such misconceptions are not only prevalent, but also difficult to revise (Sinaïtis & Broughton, 2011). Revision efforts are important because incorrect knowledge is detrimental to the development of accurate understanding and decision-making (Carey, 1985).

Revision efforts in educational settings have often used refutation texts (Kendeou, Braašch, & Bräten, 2016). These texts promote knowledge revision by explicitly acknowledging common misconceptions, directly refuting them, and providing explanations (Hynd, 2001). To date, a great deal of evidence supports the utility of refutation texts in facilitating knowledge revision (e.g., Ariasi & Mason, 2011; Broughton, Sinatra, & Reynolds, 2010; Chinn & Brewer, 1993; Guzzetti, Snyder, Glass, & Gamas, 1993; Hynd & Alvermann, 1986; Kendeou et al., 2016; Sinatra & Broughton, 2011; Tippett, 2010). However, further research is needed to identify mechanisms that help explain how such texts exert their influence (Rusanen, 2014).

Readers’ understanding of material after reading refutation texts is one potential source to determine whether knowledge revision has taken place. Such understanding is often evaluated with the use of researcher developed post-tests; indeed, higher post-test scores after reading refutation texts compared to non-refutation texts have been proposed to be indicative of learning from text and knowledge revision (Braašch, Goldman, & Wiley, 2013; Kendeou & van den Broek, 2007; Mason, Zaccoletti, Carretti, Scrimin, & Diakidoy, 2018). These post-tests often require readers to explain scientific concepts in their own words. Above and beyond their accuracy, the quality of such responses can provide information about readers’ reasoning and knowledge.

The aim of the present study was to examine the effect of refutation texts on generating explanations and explore individual differences in the quality of explanations. We draw on the extant literatures of refutation text and explanation (e.g., Chin & Brown, 2000; Kendeou et al., 2014) to understand the role of refutation texts in the generation of explanations, and determine the extent to which explanation quality can provide information about readers’ reasoning and understanding of the text.

1.1. Refutation texts and knowledge revision

One framework that has been proposed to account for knowledge
revision in the context of reading refutation texts is the Knowledge Revision Components framework (KReC; Kendeou & O’Brien, 2014). The KReC framework consists of five principles that outline the assumptions and conditions necessary for knowledge revision during reading: encoding, passive activation, co-activation, integration, and competing activation. According to the encoding principle, readers cannot simply erase information that has been encoded into long-term memory. According to the passive activation principle, this previously-encoded information can be passively reactivated during reading. These two assumptions can explain why previously acquired incorrect knowledge has the potential to be reactivated and disrupt reading, thus hindering the learning of new, related information. KReC also outlines three conditions that can facilitate revision. The co-activation principle posits that for the revision to take place, new correct information must make contact with previous, incorrect information in working memory. The integration principle posits that the result of co-activation will be the integration of new information with previous information in readers’ memory. Finally, the competing activation determines the success of revision as follows: as the recently encoded correct information gets highly integrated with additional information to support it, it draws most of the activation to itself and away from the previously acquired incorrect information. When this competition results in the reduction of the activation of the previously acquired incorrect information, knowledge revision has taken place.

The revision mechanism proposed in KReC is supported by several studies that used refutation texts. For example, evidence for the co-activation principle has been obtained in studies using a reading time paradigm, suggesting readers experience interference and slow down during the reading of correct information in refutation texts (Ariasi & Mason, 2011; Kendeou et al., 2014; Kendeou & van den Broek, 2007). Some level of knowledge revision occurs if both conceptions are co-activated and the reader experiences cognitive conflict by recognizing the conflicting information (McCrudden & Kendeou, 2014; van den Broek & Kendeou, 2008). Additionally, when readers integrate the explanation component of the refutation text, their representation of the information becomes focused on the accurate information and reduces the impact of inaccurate information (Kendeou et al., 2014; Kendeou, Butterfuss, Van Boekel, & O’Brien, 2017; Kendeou, Smith, & O’Brien, 2013; Van Boekel, Lassonde, O’Brien, & Kendeou, 2017). Indeed, Kendeou et al. (2014) provided evidence that the explanation component in the refutation text’s structure ultimately determines knowledge revision. It follows that readers’ understanding of the explanations provided in the refutation texts is therefore an important factor to consider. We turn to this issue next.

1.2. Importance and characteristics of explanations

When readers construct their own explanations, they may integrate new knowledge within their existing knowledge base and understanding of the world (Brown & Campione, 1990; Chi, De Leeuw, Chiu, & LaVancher, 1994). Asking readers to explain information they read challenges them to evaluate and elaborate on their understanding, which can positively contribute to their learning of the targeted subject matter. Indeed, the ability to generate responses that are accurate and of good quality to explain phenomena is a critical part of learning (Chin & Brown, 2000; Sandoval & Reiser, 2004). This is particularly true in science given that understanding and explaining the ‘how’ and the ‘why’ are overarching objectives of the field. Explanations are not only crucial to our understanding of the world, but they are also the way in which we exchange knowledge with others (Lombrozo, 2006).

Previous literature has determined several quality dimensions of explanations that can be either beneficial or harmful to learning (Callanan & Oakes, 1992; Daghet, 1995; Jarvie & Agassi, 1967; Rips, 2002; Thagard, 1992; for a review, see Vlach & Noll, 2016). These dimensions involve prior knowledge, understanding relationships between concepts, and the ability to make connections beyond the texts themselves. The dimensions positively related to learning include mentioning trustworthiness (e.g., Brätén, Stramsø, & Britt, 2009), using analogies correctly (e.g., Vendetti, Matlen, Richland, & Bunge, 2015), describing causal relationships (e.g., Fender & Crowley, 2007; Sandoval, 2003), utilizing prior knowledge (e.g., Asoko, 2002; Fender & Crowley, 2007; Kendeou & van den Broek, 2005), and describing physical characteristics (e.g., Asoko, 2002; Szechtler & Carey, 2009). The dimensions negatively related to learning include using personalisation (e.g., Kallery & Psillos, 2004; Legare, Lane, & Evans, 2013), describing religious, magic, or mythical forces (e.g., Browne & Woolley, 2004; Canfield & Ganea, 2014), providing circular reasoning (e.g., Baum, Danovitch, & Keil, 2008; Keil, 2006; Kendeou & van den Broek, 2005; Pritchard, 1990; Weisberg, Keil, Goodstein, Rawson, & Gray, 2008), using unnecessary descriptors (e.g., Kendeou & van den Broek, 2005; Mayer & Jackson, 2005; Pritchard, 1990), and using uncertainty descriptors (e.g., Kendeou & van den Broek, 2005; Pritchard, 1990). Utilization of these dimensions indicates one’s lack of understanding regarding the evidence that describes causal relationships and reasoning of scientific concepts.

To summarize, the ability to generate responses that are accurate and of good quality to explain concepts is critical for knowledge acquisition and revision (e.g., Chin & Brown, 2000). In turn, examining characteristics of readers’ explanations after reading refutation and non-refutation texts can provide a useful tool for revealing their knowledge and understanding of science texts. Even though there is existing evidence supporting the utility of refutation texts in facilitating knowledge revision (Guzzetti et al., 1993; Sinatra & Broughton, 2011; Tippett, 2010), further research can provide valuable insights into the characteristics of explanations readers generate as well as their reasoning after reading these texts (Brown & Campione, 1990).

1.3. Present study

The aim of the present study was to investigate the effects of refutation texts on generating explanations after reading and explore individual differences in the use of quality indicators in explanations. Our work focused on two research questions. First, we wanted to understand the extent to which explanations produced after completing reading are influenced by text structure, namely refutation and non-refutation texts. The current analysis focuses on an in-depth coding of secondary data that consisted of post-test responses readers generated after reading refutation and non-refutation texts targeting common misconceptions (Kendeou, Butterfuss, Kim, & Van Boekel, in press). We hypothesized that after completing reading, readers in the refutation condition would provide explanations with high-quality indicators for target concepts compared to readers in the non-refutation condition. Second, we explored individual differences in the quality and accuracy of explanations generated after reading refutation and non-refutation texts. We hypothesized distinct profiles would emerge using the accuracy and quality dimensions.

We approached these questions with both a variable-centered and a person-centered approach. A variable-centered approach allows for the analysis of differences between distinct groups of individuals (Block, 1971; Trevors, Kendeou, Brätén, & Brasch, 2017). In the context of this study, a variable-centered approach allowed for the analysis of differences in quality and accuracy variables between the readers’ responses in the refutation and non-refutation conditions. A person-centered approach examines individuals to determine the presence of individual differences and form similar clusters (Bergman, Magnusson, & El-Khoury, 2003; Block, 1971; Trevors et al., 2017). This approach allows for seeing how individuals differ and how variables are related. In the context of this study, a person-centered approach allowed for the examination of individual differences in the quality of explanations and their relation to learning from texts.
2. Method

2.1. Participants

The post-test data of 75 undergraduate students (44 female; M_{age} = 19.76, SD_{age} = 3.50, Range_{age}: 18–38 years) enrolled in introductory psychology courses were used in the current study. Participants received partial course credit for their participation.

2.2. Materials

2.2.1. Texts

Participants read 10 narrative-informational texts (Duke, 2000), each addressing a common misconception that had been identified as highly prevalent in the population from which the current sample was drawn (e.g., Van Boekel et al., 2017). Narrative-informational texts were chosen because they allow opportunities to refute and explain common misconceptions in a friendlier context for the reader to comprehend (Kendeou et al., 2014). Appendix A lists the 10 common misconceptions. Appendix B contains sample texts.

The refutation texts included a refutation-plus-ellation section that explicitly stated and refuted the target misconception. The refutation was immediately followed by an explanation that provided extra information supporting the refutation. The non-refutation text included a non-refutation-plus-non-ellation section followed by the same structure of the refutation section. However, the text in this section progressed the story, making no mention of either the misconception or the refutation. Both texts included a correct outcome (target) sentence that stated the correct information. Each passage ended with a comprehension question that did not address information concerning the misconception but was included to ensure understanding.

Two material sets were constructed and each contained 10 texts, half in each of the two experimental conditions. Across the two sets, each text occurred once in each condition. Participants were randomly assigned to a set.

2.2.2. Post-test

The test included 10 two-tiered questions (Tan, Goh, & Chia, 2001; Tregast, 1988) corresponding to the 10 misconceptions targeted in the texts. The first tier was a True/False question, followed by the second tier asking readers to provide an explanation for their True/False response. The two-tiered questions afforded the opportunity to assess the readers’ knowledge. For instance, an incorrect answer to the first-tier item and an incorrect explanation illustrates the existence of incorrect knowledge. A correct answer to the first-tier item and an incorrect explanation also illustrates the presence of incorrect knowledge. Conversely, a correct answer to the first-tier item and a correct explanation illustrates correct knowledge.

2.3. Procedure

All procedures, including the informed consent and the recruitment of participants, were reviewed and approved by the University’s Institutional Review Board (IRB). Participants were tested individually in a single session. They were informed they were going to read texts for general understanding. Approximately half of the participants were asked to think-aloud as they read, and the remaining half were asked to read silently (for detailed methods see Kendeou et al., in press). When the participants finished reading all texts, they were asked to complete the post-test. To address the aims of this study, we performed an elaborate coding of the post-test responses.

2.3.1. Coding post-test responses

Each written response readers generated on the post-test was coded with respect to accuracy and quality dimensions. The unit of analysis was the participant’s entire written response. The dimensions for each response were coded independently of one another in terms of presence (code = 1) or absence (code = 0) in each response. Two raters independently coded all post-tests. To determine inter-rater reliability, 20% of the post-tests were compared. Disagreements were resolved through discussion. Inter-rater reliability was high (96.9%).

Drawing on the extant literature, 13 coding dimensions were selected that reflected the accuracy and quality characteristics of a response. Table 1 lists each accuracy dimension, the corresponding definition, and an example from the written post-test responses. Table 2 lists each quality dimension, the corresponding definition, and an example from the written post-test responses.

Two accuracy dimensions evaluated whether each response was an explanation (code = 1) or not (code = 0), and the accuracy of the response was in comparison to the information in the text (Chin & Brown, 2000; Lombrozo, 2006). Codes for the scientific accuracy of the response quantified the level at which the reader understood the correct reason for why the belief statement was true or false, either based on text information or prior knowledge (Kendeou & van den Broek, 2005; Pritchard, 1990; Ruiz-Primo, Li, Tsai, & Schneider, 2010). The response was coded as accurate (2 points) or inaccurate (0 points).

The quality of readers’ responses was assessed using 11 dimensions. Previous literature has determined are either beneficial or harmful to learning science (Callanan & Oakes, 1992; Dagher, 1995; Jarvie & Agassi, 1967; Rips, 2002; Thagard, 1992; Vlach & Noll, 2016).

3. Results

Before performing any analyses to address the main research questions, we examined if there were differences between the think-aloud and silent reading groups in each of the dependent variables. A series of mixed ANOVAs showed no differences in any of the dependent variables (all Fs < 1). Thus, data from both groups were combined for subsequent analyses.

3.1. Research question 1

3.1.1. Quality

First, we wanted to understand the extent to which explanations generated after reading were influenced by text structure, namely refutation and non-refutation results. Results showed, on average, significantly more explanations were generated in the refutation (M = 4.60, SD = 0.90) than the non-refutation condition (M = 3.81, SD = 1.21), t(74) = 5.62, p < .001, d = 0.74. Also, accuracy of responses was significantly higher in the refutation (M = 6.96, SD = 3.07) than the non-refutation condition (M = 3.09, SD = 2.62), t(74) = 9.03, p < .001, d = 1.36. Causality was significantly higher in the refutation than the non-refutation condition. Circularly and
uncertainty descriptors were significantly higher in the non-refutation condition. No other effects were significant. Table 3 depicts differences by condition.

### 3.2. Research question 2

#### 3.2.1. Profiles analysis

Second, we wanted to explore if there were individual differences in quality and accuracy of explanations generated after reading refutation and non-refutation texts. In doing so, we used cluster analysis. To determine the appropriate number of clusters, we evaluated two-, three-, and four-cluster solutions for distinct profiles with adequate numbers of participants (Pastor, 2010). The dimensions of analogy and religious/magical were not included in the analyses since they had zero frequency. The ANOVA results for the two-cluster solution did not significantly differentiate among the nine quality dimensions (all ps > .05) for trust, mistrust, causality, prior knowledge, and circularity. The ANOVA results for the three-cluster solution showed acceptable differentiation among the nine quality dimensions (all ps < .10). The ANOVA results for the four-cluster solution did not significantly differentiate among the nine quality dimensions (all ps > .05). Thus, the three-cluster solution was determined to produce the most distinct profiles. Profiles are presented in Table 4, along with the results from

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Quality: dimensions, definitions, and examples from post-test explanations.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension</td>
<td>Definition</td>
</tr>
<tr>
<td>Trust</td>
<td>The response addressed trust of the text.</td>
</tr>
<tr>
<td>Analogy</td>
<td>The response used a “like” or “as” comparison.</td>
</tr>
<tr>
<td>Causal relationships</td>
<td>The response described a cause and effect relationship to explain the belief statement.</td>
</tr>
<tr>
<td>Prior knowledge</td>
<td>The response mentioned information the reader recognized from the text or reader's prior experiences.</td>
</tr>
<tr>
<td>Physical characteristics</td>
<td>The response described physical traits or objects that were not described in the belief statement.</td>
</tr>
<tr>
<td>Mistrust</td>
<td>The response addressed mistrust of the text.</td>
</tr>
<tr>
<td>Personification</td>
<td>The response gave a human quality to an object or idea that was not human.</td>
</tr>
<tr>
<td>Religion/magic</td>
<td>The response included religious or magic information.</td>
</tr>
<tr>
<td>Circularity</td>
<td>The response was repetitive.</td>
</tr>
<tr>
<td>Uncertainty descriptors</td>
<td>The response included phrases showing uncertainty about the information.</td>
</tr>
</tbody>
</table>

#### 3.2.2. Clusters

Cluster 1 was identified as coherence-building readers because their explanations reflected several high-quality indicators. Specifically, Cluster 1 (n = 46) used significantly more causality (p < .001) and prior knowledge (p < .001) than Cluster 2. Cluster 1, also used significantly less trust (p = .019), mistrust (p = .034), circularity (p < .001), and uncertainty descriptors (p < .001) than Clusters 2 and 3. There were no significant differences between Clusters 1 and 2, for physical characteristics, personification, and unnecessary descriptors (all ps > .05).

Cluster 2 was identified as non-coherence-building readers because their explanations reflected several low-quality indicators. In turn, Cluster 2 (n = 26) used significantly less causality (p = .013), physical characteristics (p = .002), personification (p = .001), unnecessary descriptors (p = .001), and uncertainty descriptors (p = .002) than Cluster 3. There were no significant differences between Cluster 2 and Cluster 3 for trust, mistrust, prior knowledge, and circularity (all
ps > .05).

Finally, Cluster 3 was identified as promiscuous readers because their explanations included both high- and low-quality indicators. This cluster included three exceptional cases. Cluster 3 (n = 3) used significantly more physical characteristics (p = .003), personification (p = .001), unnecessary descriptors (p < .001), and uncertainty descriptors (p < .001) than Cluster 1. There were no significant differences between Cluster 3 and Cluster 1 for trust, mistrust, causality, prior knowledge, and circularity (all ps > .05).

### 3.2.3. Cluster validation

To validate the three clusters, additional analyses were performed. First, a Kruskal-Wallis analysis was conducted with cluster membership as the independent variable and accuracy scores, a variable not used in the cluster analysis, as the dependent variable. Results showed a significant difference between the three clusters (p < .001). Follow-up post-hoc comparisons using Mann-Whitney showed Cluster 1 had higher accuracy scores than Cluster 2 (p < .001) and Cluster 2 had higher accuracy scores than Cluster 3 (p < .05). Cluster 1 also had higher accuracy scores than Cluster 3 (p < .001). These differences are consistent with the characteristics of each cluster.

Second, a Kruskal-Wallis analysis was conducted with cluster membership as the independent variable and explanation scores, another variable not used in the cluster analysis, as the dependent variable. The explanation score refers to the dimension which assessed if the response was an explanation. Results showed a significant difference between the three clusters (p < .001). Follow-up by post-hoc comparisons using Mann-Whitney showed Cluster 1 generated more explanations than Cluster 2 (p = .045) and Cluster 2 generated more explanations than Cluster 3 (p = .017).

### 4. Discussion

The aim of the present study was to examine the effects of refutation texts on explanation quality and explore potential individual differences. To achieve this aim we sought to answer two specific research questions. First, we examined the extent to which explanations produced after completing reading were influenced by text structure, namely refutation and non-refutation texts. We hypothesized that after reading was completed, readers in the refutation condition would provide explanations with high-quality indicators for target concepts compared to readers in the non-refutation condition. Indeed, readers in the refutation condition provided significantly more explanations and used causality, a high-quality indicator, more than readers in the non-refutation condition. Conversely, the explanations from readers in the non-refutation condition contained significantly more circular phrases and uncertainty descriptors than readers in the refutation condition.

Second, we explored whether there were individual differences in quality and accuracy of explanations generated after reading refutation and non-refutation texts. We hypothesized distinct profiles would emerge using the quality dimensions and accuracy. Indeed, a three-cluster solution demonstrated three types of readers: coherence-building readers, non-coherence building readers, and cases of promiscuous readers, each negotiating the reading task differently and establishing different levels of accuracy and understanding. The coherence-building readers were distinguished by their high use of causality and high post-test scores suggesting success in learning from texts. The non-coherence building readers were distinguished by their high use of circularity and low post-test scores suggesting less success in learning. The promiscuous readers were special cases distinguished by their use of both low-quality (personification, unnecessary descriptors, uncertainty descriptors) and high-quality indicators (physical characteristics) and the lowest post-test scores, suggesting low levels of learning.

There is precedent in the literature identifying similar clusters or profiles of readers across development. Thus, these profiles may represent relatively stable individual differences (Gernsbacher, 1997). For example, Bohn-Gettler and Kendeou (2014) identified clusters of coherence-building and non-coherence-building processes in a sample of adult readers. The extent to which readers engaged in coherence or non-coherence building depended on individual differences in working memory capacity. Importantly, coherence-building processes were associated with high levels of comprehension, whereas non-coherence building processes were associated with low levels of comprehension. This pattern of results is consistent with the coherence and non-coherence building readers identified in the present study. Related, McMaster and colleagues (McMaster et al., 2012; Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007) identified several clusters of young readers in grade 4, including a group of struggling readers who appeared to engage in activities much as good readers did. However, these activities were frequently unsuccessful and also drew on inappropriate knowledge. This pattern of results is consistent with the promiscuous readers identified in the present study, who frequently used both high- and low-quality indicators in their explanations, resulting in low overall post-test performance.

Taken together, the results suggest readers in the refutation condition reached a better understanding of causal relationships, and generated more accurate responses with less circularity and uncertainty, than the non-refutation condition. One explanation for these findings is that refutation texts may have helped readers to break the “cycle of circularity”; readers may have gained an understanding of the direct causal pathways of the conception and thus relied less on circular misconceptions. Consequently, the refutation texts then supported the generation of explanations with quality dimensions, which in turn facilitated comprehension, articulation, and perhaps also knowledge revision. Indeed, this study suggests refutation texts may engender shifts in the directionality of thought, from circular to causal pathways, which is an important milestone in scientific understanding (McNeill, Lizotte, Krajcik, & Marx, 2006; Zohar & Nemet, 2002). Future research should target these thought processes as they may be the foundation for knowledge revision.

It is important to note that the refutation texts in the present study addressed misconceptions that were conceived at the individual belief level (Chi, 2008). Previous research has shown that such misconceptions are possible to revise after reading refutation texts (Braasch et al., 2013; Kendeou & van den Broek, 2007; Mason et al., 2018), whereas other types of misconceptions that are more complex are less likely to be revised (Chi, 2013). Even though we cannot be certain that knowledge revision has taken place in this study given the absence of a pre-test, it is very likely that it did. Indeed, if we integrate both process (e.g., think-aloud protocols, reading times; Kendeou et al., in press) and outcome measures, there is convergence for knowledge revision. The present analysis, however, focused exclusively on outcome measures, thus we refrain from making strong claims about revision.

From a methodological point of view, the present study demonstrates how variable-centered and person-centered approaches can be integrated to produce new insights into readers’ learning from text. A variable-centered approach allowed for the identifications of differences between explanations produced as a result of reading refutation and non-refutation texts, highlighting the importance of text structure. A person-centered approach allowed for the identification of individual differences in the quality of the explanations produced, highlighting the presence of clusters or profiles of readers (Bohn-Gettler & Kendeou, 2014; McMaster et al., 2012; Rapp et al., 2007; Trevors et al., 2017). Thus, the combination of both variable-centered and person-centered approaches produced new insights into how text structure influences learning from text and how individuals respond to different text structures.

Despite these contributions, this study also has its limitations. As we noted above, one limitation is the lack of a pre-test of misconceptions for the specific sample. Even though previous research found that the targeted misconceptions are prevalent in the population from which the current sample has been drawn (e.g., Kendeou et al., 2014; Van Boekel
et al., 2017), it is very likely that there was variability in readers' pre-existing beliefs. In future research, a pre-test should be incorporated to confirm the presence and strength of misconceptions prior to engaging with refutation and non-refutation texts. The challenge of including a pre-test is determining how it could be incorporated in this context without drawing attention to the target misconceptions or activating them before reading and thus potentially influencing post-test outcomes. Future research should also consider the inclusion of a delayed post-test measure to provide insight into whether readers retained the correct knowledge and maintained the reported learning gains long-term. Additionally, continued work is necessary to understand the negative relation between trust as a quality indicator and accuracy. This was a surprising finding and may be due, in part, to the low frequency of this dimension in student responses.

4.1. Conclusions

The current results add to the growing body of work showing that refutation texts are a powerful tool for learning (Ariasi & Mason, 2011; Broughton et al., 2016; Chinn & Brewer, 1993; Danielson, Sinatra, & Kendeou, 2016; Guzzetti et al., 1993; Hynd & Alvermann, 1986; Kendeou et al., 2016; Sinatra & Broughton, 2011; Tippett, 2010). This work also demonstrates that examining characteristics of explanations is a useful method for revealing how learners' reasoning influences their understanding of and learning from texts. In this context, refutation texts may facilitate shifts from circular to causal reasoning pathways that, in turn, may be foundational for knowledge revision.

Appendix A

A.1. Targeted common misconceptions

1. Seasons are caused by the Earth being closer to the Sun in the summer than in the winter.
2. Meteors that land on Earth are hot.
3. Napoleon Bonaparte was very short.
4. Chameleons change color to match their surroundings.
5. People only use 10% of their brains.
6. If you drop two balls of the same size but one weighs twice as much, the heavier ball will hit the ground first.
7. Reading in dim light causes nearsightedness (myopia).
8. Eating turkey makes people especially drowsy.
9. Ostriches bury their heads in the sand.
10. Lightning never strikes the same place twice.

Appendix B

B.1. Sample refutation text

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 she 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

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.

Elaboration

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.

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. She listened carefully as Jerry assured everyone that...
texts: The role of co-activation in confronting misconceptions. Applied Cognitive Psychology, 22, 335–351.