

Refutations in science texts lead to hypercorrection of misconceptions held with high confidence

Citation for published version (APA):

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. <https://doi.org/10.1016/j.cedpsych.2015.04.003>

Document status and date:

Published: 01/07/2015

DOI:

[10.1016/j.cedpsych.2015.04.003](https://doi.org/10.1016/j.cedpsych.2015.04.003)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

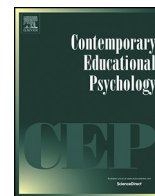
providing details and we will investigate your claim.



ELSEVIER

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Refutations in science texts lead to hypercorrection of misconceptions held with high confidence



Mariëtte H. van Loon ^{a,b,*}, John Dunlosky ^c, Tamara van Gog ^d,
Jeroen J.G. van Merriënboer ^b, Anique B.H. de Bruin ^b

^a Department of Developmental Psychology and Swiss Graduate School for Cognition, Learning and Memory, University of Bern, Fabrikstrasse 8, CH-3012 Bern, Switzerland

^b Department of Educational Development & Research and Graduate School of Health Professions Education, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

^c Department of Psychology, Kent State University, P.O. Box 5190, Kent, OH 44242-0001, USA

^d Department of Psychology, Erasmus University Rotterdam, P.O. Box 1783, 3000 DR, Rotterdam, The Netherlands

ARTICLE INFO

Article history:

Available online 14 April 2015

Keywords:

Reading comprehension
Confidence
Misconceptions
Correction
Adolescence

ABSTRACT

Misconceptions about science are often not corrected during study when they are held with high confidence. However, when corrective feedback co-activates a misconception together with the correct conception, this feedback may surprise the learner and draw attention, especially when the misconceptions are held with high confidence. Therefore, high-confidence misconceptions might be more likely to be corrected than low-confidence misconceptions. The present study investigates whether this hypercorrection effect occurs when students read science texts. Effects of two text formats were compared: Standard texts that presented factual information, and refutation texts that explicitly addressed misconceptions and refuted them before presenting factual information. Eighth grade adolescents ($N = 114$) took a pre-reading test that included 16 common misconceptions about science concepts, rated their confidence in correctness of their response to the pre-reading questions, read 16 texts about the science concepts, and finally took a post-test which included both true/false and open-ended test questions. Analyses of post-test responses show that reading refutation texts causes hypercorrection: Learners more often corrected high-confidence misconceptions after reading refutation texts than after reading standard texts, whereas low-confidence misconceptions did not benefit from reading refutation texts. These outcomes suggest that people are more surprised when they find out a confidently held misconception is incorrect, which may encourage them to pay more attention to the feedback and the refutation. Moreover, correction of high-confidence misconceptions was more apparent on the true/false test responses than on the open-ended test, suggesting that additional interventions may be needed to improve learners' accommodation of the correct information.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Adolescents in secondary education are increasingly expected to engage in self-regulated learning in many subjects, and science is no exception (Alvermann, 2002). Self-regulated learning of science texts can be challenging, especially when learners are required to understand novel concepts presented in the texts (Otero, Leon, & Graesser, 2002; Sinatra & Broughton, 2011). For instance, learners have difficulty understanding concepts explained in texts, such as airflow (Braasch, Goldman, & Wiley, 2013), motion (McCloskey,

1983), or the nature of light (Mason, Gava, & Boldrin, 2008). Furthermore, learners often cannot accurately monitor whether they correctly understand the information presented in studied texts (Thiede, Griffin, Wiley, & Redford, 2009). When monitoring of text learning is inaccurate, self-regulated learning is not likely to be well adapted to the current level of understanding. Because learners cannot strategically decide which passages are not yet well learned, they may stop studying prematurely (Metcalfe & Finn, 2008; Rawson & Dunlosky, 2013; Son & Metcalfe, 2000; Thiede, Anderson, & Theriault, 2003).

According to Piaget (1964), learning is a transformative process that requires modification of existing cognitive schemata through assimilation and accommodation. When a science text contains information that extends or deepens existing knowledge, learners need to assimilate the novel information into their existing cognitive schema. Alternatively, when the new information is not compatible

* Corresponding author. Department of Developmental Psychology and Swiss Graduate School for Cognition, Learning and Memory, University of Bern, Fabrikstrasse 8, CH-3012 Bern, Switzerland. Fax: +41 31 631 82 12.

E-mail address: mariette.vanloon@psy.unibe.ch (M.H. van Loon).

with learners' cognitive schema, and conflicts with the information in their knowledge base, this means that they need to change the organization of their cognitive schema by accommodating the novel information (Kendeou & O'Brien, 2014; Piaget, 1964).

One major obstacle to learning from science texts is learners' misconceptions, which are often based on naive ideas about concepts that differ from the accepted scientific facts. For instance, Schneps (1989) reported that many university students hold the misconception that seasonal changes are caused by fluctuations in the distance between the Earth and the Sun. Instead, seasonal changes are caused by the tilted position of the Earth. Such misconceptions about scientific concepts are prevalent and resistant to change, and can therefore hinder accommodation of accurate knowledge (Alvermann & Hynd, 1989; Braasch et al., 2013; Hammer, 1996; Kendeou & van den Broek, 2005).

1.1. Confidence and hypercorrection of misconceptions

Schneps (1989) found that students were very confident in their misconceptions, suggesting that these are firmly established. The relation between confidence in misconceptions and correction of misconceptions seems to be complex. On the one hand, high confidence in prior knowledge negatively impacts learners' success in accommodating contrasting information (Pintrich, Marx, & Boyle, 1993). Misconceptions that learners state with high confidence are resistant to change (Dole & Sinatra, 1998; Ecker, Lewandowsky, & Tang, 2010). Presumably, high-confidence misconceptions are harder to correct than the ones that are held with lower confidence, because high-confidence misconceptions are more strongly represented in memory (Ecker, Lewandowsky, Swire, & Chang, 2011). On the other hand, unless learners are given feedback, correction of misconceptions is unlikely (Metcalfe, Butterfield, Habeck, & Stern, 2012); and contrary to intuition, feedback may help most when learners are highly confident in a misconception (Butterfield & Metcalfe, 2001; Cordova, Sinatra, Jones, Taasobshirazi, & Lombardi, 2014). In particular, when learners receive explicit feedback that they hold a misconception, the misconceptions that are held with high confidence are more likely to be corrected than misconceptions that are held with less confidence, a finding called the *hypercorrection effect* (Butterfield & Metcalfe, 2001).

In prior research on hypercorrection, participants answered general information questions such as: What is the name of the river that runs through Rome? After giving an answer, they rated their confidence in the response. When they answered questions incorrectly, they received feedback immediately in the form of the correct answer (Butterfield & Metcalfe, 2001; Metcalfe & Finn, 2012). On a subsequent test, it appeared that high-confidence errors were more often corrected than low-confidence errors. Presumably, when people confidently make an incorrect response, the correcting feedback surprises them and draws attention to the feedback (e.g., Butler, Fazio, & Marsh, 2011; Metcalfe et al., 2012). People are more surprised when feedback contradicts their high-confidence misconceptions than their low-confidence ones, so they might pay extra attention to the correct information, which in turn leads to the hypercorrection effect (Butterfield & Metcalfe, 2006; Metcalfe & Finn, 2011, 2012).

1.2. Correcting misconceptions with refutation texts

Whether the hypercorrection effect will also occur when reading texts about science concepts, instead of learning short answers to general information questions (which were the materials used in prior research), is not yet known. When reading science texts, simply giving feedback that an answer is not correct is usually insufficient for conceptual change to occur (for a review, refer to Guzzetti, Snyder, Glass, & Gamas, 1993). In most educational science text

books, standard text formats are used that emphasize and explain currently accepted scientific explanations of concepts (Tippett, 2010). When reading standard texts, learners might have difficulties with noticing the discrepancy between their misconceptions and the provided explanation for the correct conception. Therefore, standard text formats may not be effective for correcting misconceptions, because explanations in the studied text do not undermine the perceived truth of learners' science misconceptions (Braasch et al., 2013; Diakidoy, Kendeou, & Ioannides, 2003).

When studying science texts, providing explicit refutations of misconceptions seems to increase the likelihood of correction (Guzzetti et al., 1993). To correct misconceptions, learners need to co-activate these together with the correct concepts to become aware that the two conflict with each other (van den Broek & Kendeou, 2008). A special text format, called a refutation text, has been designed to support such co-activation. *Refutation texts* always contain two components in addition to a standard text: The statement of the misconception and explicit refutation of this misconception (Braasch et al., 2013; Tippett, 2010). When explicitly refuting the misconception and co-activating it with the correct concept, the inconsistency between a learner's misconceptions and the contrasting information is likely to lead to an experience of cognitive dissonance (Guzzetti et al., 1993). Because reading refutation texts increases awareness of conflicts between a learner's prior beliefs and new information, this text format helps the learner to correct misconceptions.

With this study, we aim to address the question: Does reading refutation texts also lead to hypercorrection for high-confidence misconceptions? When high-confidence misconceptions are activated and then refuted in the text, learners might be surprised and more fully attend to the correct conception than when it is presented using standard texts (Broughton, Sinatra, & Reynolds, 2010). Hence, a hypercorrection effect might occur when learners read refutation texts on topics for which they hold high-confidence misconceptions. This study investigates both the processes of 'outdating', which is negating the misconception, and 'updating', which is accommodating the correct alternative in the memory representation (Kendeou, Smith, & O'Brien, 2013). Moreover, it is important to investigate whether hypercorrection (if it occurs) would remain after some delay. In educational settings, learners are typically not tested immediately after studying texts but after a delay, and hence they need to retain the studied information in long-term memory. However, almost all research on hypercorrection has used a delay of only a few minutes between the presentation of feedback and the final test, although outcomes from two studies suggest that the hypercorrection effect persists after one week (Butler et al., 2011; Butterfield & Mangels, 2003). We investigated whether reading refutation texts, containing explicit feedback about how their misconceptions are incorrect, has beneficial effects on adolescents' correction (i.e., outdating and updating) of high-confidence science misconceptions on a delayed test.

Accordingly, the main purpose of this study was to test the refutation-causes-hypercorrection (RCH) hypothesis, which predicts that reading refutation texts leads to enhanced correction of high-confidence misconceptions in comparison to reading standard texts; for correction of low-confidence misconceptions reading refutation texts is hypothesized to have less of an effect (Hypothesis 1). Moreover, consistent with Sanchez and Garcia-Rodicio (2013), we expected that reading refutation texts would specifically focus learners on correction of misconceptions, so that text format would not affect learning of details (Hypothesis 2). In the present study, we investigated eighth graders' correction of misconceptions. Half of the learners studied refutation texts and the other half studied standard texts. A post-test administered one week after studying the science texts assessed whether learners outdated their misconceptions and were able to update their misconceptions with the

correct information. Furthermore, detail questions were asked about information in the texts that was not directly related to the misconceptions.

1.3. Monitoring and regulation of correction of misconceptions

The majority of research on effects of refutation texts has focused on improvement of learners' performance *per se* (i.e., cognition). Even though this focus is important, it is also relevant to investigate how performance improvements come about. What cognitive and metacognitive processes during self-regulated study of refutation science texts result in correction of misconceptions? The current study is grounded in the framework presented by Nelson and Narens (1990), which emphasizes the relation between cognition and metacognition during learning. When studying texts, learners need to monitor their learning and regulate the learning process to reduce the discrepancy between intermediate states of learning and their overall learning goals (Nelson & Narens, 1990; Thiede & Dunlosky, 1999).

In a self-regulated learning setting, performance of a learner is affected by the accuracy of their monitoring of learning. A learner who monitors that studied information is not yet sufficiently understood should decide to continue studying instead of dropping the materials from further learning. Only when students accurately monitor their learning can they engage in effective self-regulated learning (Dunlosky, Hartwig, Rawson, & Lipko, 2011; Dunlosky & Thiede, 2013; Metcalfe & Finn, 2008). Monitoring of text learning

would likely be inaccurate when learners hold misconceptions, because they often perceive this knowledge as correct when, in fact, it is entirely inaccurate (Van Loon, de Bruin, van Gog, & van Merriënboer, 2013). When learners do not accurately monitor their understanding of studied texts, their confidence in the accuracy of their misconception can lead to premature termination of study (Van Loon et al., 2013). Even though refutation texts have been demonstrated to improve learners' performance (in terms of the cognitive level from Nelson and Narens' framework), it is unknown whether reading refutation texts can improve monitoring accuracy and regulation (the metacognitive level) when reading texts about misconceptions. Therefore, a final question of this research is: Could refutation texts potentially improve learners' ability to accurately monitor their understanding of studied texts and help them to make adaptive restudy selections?

1.4. Research questions and hypotheses

Fig. 1 shows a graphical overview of the present study in relation to the framework of Nelson and Narens (1990). As shown in Fig. 1, both text format and subjective confidence in misconceptions are expected to affect learners' performance (their cognitive level). Reading refutation texts is hypothesized to lead to enhanced correction of high-confidence misconceptions in comparison to reading standard texts; for correction of low-confidence misconceptions, reading refutation texts is hypothesized to have less of an effect (Hypothesis 1). We do not expect an effect of text format

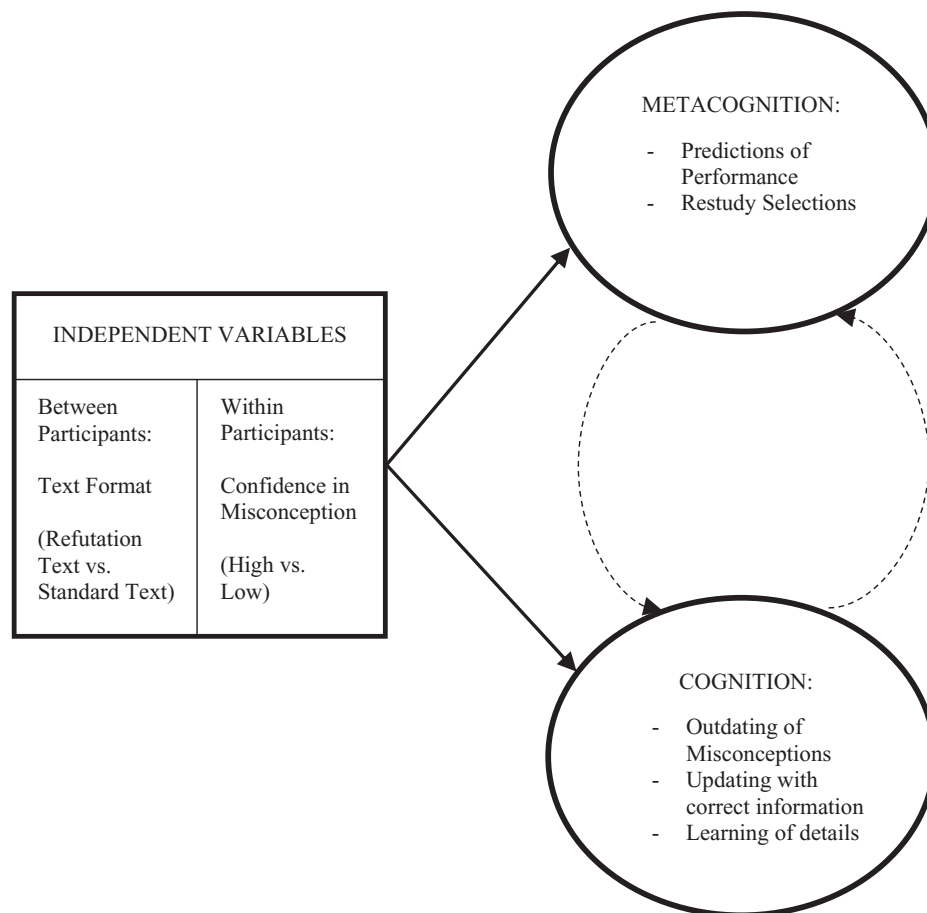


Fig. 1. Hypotheses and Open-Ended Questions. This diagram depicts the hypothesized effect of Text Format and Confidence in Misconception on cognition, and the open-ended research questions about effects on metacognition. The box represents the independent variables; ellipses represent dependent variables; solid arrows represent the hypotheses and open-ended research questions; dashed arrows represent the relation between cognition and metacognition, based on the framework by Nelson and Narens (1990).

on learning of details (Hypothesis 2). As depicted in Fig. 1, in addition to investigating the effects of text format and confidence in cognition, we address whether reading refutation texts improves learners' metacognition. Monitoring accuracy is measured by relating participants' predictions of performance (POPs) to their test performance, regulation is measured with restudy selections. We address the open questions as to whether reading refutation texts improves POPs (Question 1) and adaptive restudy selections (Question 2) in comparison to reading standard texts.

2. Method

2.1. Participants and design

Participants were 114 eighth-grade pupils (M age = 13.75 years, $SD = .51$, 70 females, 44 males) of a secondary school in The Netherlands (all participants followed education at the highest level of secondary education, with a duration of 6 years). Each participant was assigned randomly to one text format (refutation, standard). The experiment consisted of two separate sessions. During the first session, participants were tested in the computer room of their school. During the second session, which was one week after the first session, they were tested with a paper-and-pencil task. In the first session, 118 pupils participated, however, four participants did not attend the second session, the remaining participants completed both sessions; 57 participants read refutation texts and 57 read standard texts.

2.2. Materials

2.2.1. Pre-reading test and confidence ratings

Pre-reading test questions were true/false statements about common misconceptions. An example of a statement is: "The brightest stars are the largest stars." The statements were addressing common misconceptions that were identified in a pilot study beforehand. In this pilot study, 51 participants in the same age group as our sample answered true/false statements and then rated confidence on a scale from 0% to 100% (ranging from 'not sure whether my response is correct' to 'entirely sure that my response is correct'), with increasing steps of 20%. Participants were divided in two subgroups, two versions of the pilot study were used, and participants in each subgroup answered 20 of the 40 selected pilot statements. From these pilot data, 16 statements were selected (refer to Appendix A for the misconceptions addressed in this study). The statements were selected to ensure that these were common misconceptions that were often answered incorrectly, and to ensure that the responses to the statements would show sufficient variability in accuracy and confidence ratings. For eight of the pre-reading test questions the correct response was 'true', for the other eight the correct response was 'false'. In the pre-reading test, participants filled out a response for each statement (i.e., true or false) and then rated their confidence for this statement on the confidence scale, ranging from 0% to 100% (i.e., points on the scale matched 0%, 20%, 40%, 60%, 80%, and 100%). The text topics are independent from each other, so learners can hold one misconception without having a larger/smaller chance of holding another misconception. Thus, not unexpectedly, the internal consistency of across questions was low ($\alpha = .10$).

2.2.2. Study phase

The study materials consisted of 16 passages that had the same topics as the pre-reading test statements. The topics of these passages, the misconceptions, and the correct information for all topics are listed in Appendix A. The difference between the two experimental groups was the format of the text. The *standard* texts consisted of a passage that explained the correct information

(a contradiction of the common misconception) and details that were related to the topic of the text. The *refutation* texts contained, in addition to the correct information and the passage about details, a sentence that presented the common misconception and then explicitly refuted it. In the refutation text, an additional paragraph introduced the misconception (e.g., "many people believe that a camel's hump is filled with water"), refuted it (e.g., "this is not true"), and then provided the correct information (e.g., "a camel's hump consists of fat"). See Appendix B for examples of standard texts and refutation texts. The mean number of words per text was 108.88; $SD = 9.31$; range = 93–125 for standard texts; for refutation texts mean words per text was 126.00; $SD = 10.27$; range = 111 to 143.

2.2.3. Predictions of performance and restudy selections

When predicting their test results, participants saw the title of the text on the screen, accompanied by the question: How many test questions do you expect to complete correctly about the text [Title]? These POPs were provided with a mouse-click on a 6-point scale ranging from 0% to 100% (points on the scale matched 0%, 20%, 40%, 60%, 80%, and 100%).

When making restudy decisions, participants indicated which texts they would want to restudy by clicking on a grid with an 8x2 array in which each cell was filled by a title of a previously studied text. The order of the texts in this grid was randomized for every participant, 0 to 16 texts could be selected.¹

2.2.4. Post-test

The post-test contained three subtests: Open-ended test questions about misconceptions, true/false statement test questions, and detail questions. Appendix C presents examples of the test questions. Sixteen open-ended questions about misconceptions addressed key aspects of the common misconceptions that were contradicted by the study passages ($\alpha = .59$). The true/false test consisted of the same 16 true/false statements that were used for the pre-reading test ($\alpha = .66$). Further, 16 questions were asked about details from the studied texts ($\alpha = .50$).

2.3. Procedure

The two sessions took about 80 minutes of instructional time; the second session occurred one week after the first session. All participants were tested in groups of 27 to 32 in their classroom, with different conditions being present in each session to which the students were randomly assigned. In both sessions, all tasks were self-paced, and participants were encouraged to guess when they were not sure about their answer. Apart from the refutations in the texts for the refutation group, participants did not receive feedback in any phase of the experiment.

The first session took place in the computer room of the school. Before starting the task, the experimenter explained to the participants that they would take a pre-reading test and then study texts, and they were shown an example of the confidence scale. Further, they were instructed that they would need to take a test about the studied texts the following week, and that the test would contain both open-ended and multiple-choice questions. Participants were instructed to use their personal log-in number to start the task.

Once all pre-reading test statements had been answered, there was a short filler task in which participants completed some questions about their experience with scientific research. These filler questions were added to ensure that the study phase did not follow immediately after completing the pre-reading test. When reading

¹ Note that learners' restudy selections were not honored. We did not include an intervening restudy trial so as to not inadvertently influence the outcome measures that were most relevant to our main hypotheses.

the texts, participants could proceed to the next text with a mouse-click. After reading the texts, participants again completed short filler questions, which were four riddles. This filler task was given to ensure that POPs and restudy selections did not follow immediately after text study.

In session 2, which took place one week after the first session, participants were tested with pencil-and-paper tests about their memory for texts they studied in session 1 with the open-ended questions as well as the statement test with true/false questions.

2.4. Scoring

Scores on the open-ended questions about the misconception topic were scored as omission (0 credit), misconception (0 credit), other error (0 credit), partially correct (.5 credit), or fully correct (1 credit). For example, for the question concerning why some stars are brighter than others (which is related to how close they are to the earth), if the answer stated that the brighter stars are largest, it was scored as a misconception. If the participants did not provide any response, it was scored as an omission. If a participant stated, for example, that star brightness depends on the material a star is made of, this response was scored as 'other error'. If the participant responded that it had something to do with the distance, but did not explain that the brighter stars are closer to the earth, this response was scored as partially correct. Comprehension was emphasized; therefore, a response was also scored as correct when the participant responded with a paraphrase of the original text and spelling errors did not affect scoring. Two raters scored 20.1% (a total of 367 out of 1824) responses on open-ended questions and inter-rater agreement was high ($Kappa = .88$); the scores of the first rater were used for the analyses.

The detail questions were scored as omission (0 credit), error (0 credit), partially correct (.5 credit), or fully correct (1 credit). Two raters scored 20% (a total of 365 out of 1824) of all test responses on detail questions, and agreement was high ($Kappa = .93$).

3. Results

First, data relevant to evaluating the RCH-hypothesis (Hypothesis 1) and effects of text format on detail learning (Hypothesis 2) are presented. Next, analyses relevant to evaluating possible effects of text format on POPs (Question 1) and restudy selections (Question 2) are reported. Finally, for completeness, additional statistics

about reading times and correlations between the measured variables are presented.

3.1. Effects of text format on correction of high- and low-confidence misconceptions

The RCH hypothesis (Hypothesis 1) predicts that compared to standard texts, reading refutation texts leads to enhanced correction of high-confidence misconceptions, whereas there will be less of an effect of text format on correction of low-confidence misconceptions. Learners' confidence in misconceptions (ranging from 0% to 100% with increasing steps of 20%) was split into low confidence (0% and 20%); medium confidence (40% and 60%); and high confidence (80% and 100%) categories. Table 1 includes the overall test performance and the correction of misconceptions for the groups reading refutation texts and standard texts. As shown in Table 1, correct performance was higher on the true/false test and the open-ended test for the group reading refutation texts, and learners reading refutation texts had fewer misconceptions on the open-ended test.

Performance on the true/false statement test and on the open-ended test was directly relevant to evaluating the RCH hypothesis because both tests focused on the correction of misconceptions that could be related to confidence on the pre-reading test. (The detail test did not focus on misconceptions and could not involve a breakdown by confidence level, so we analyze that measure separately below.) Thus, to evaluate the RCH hypothesis, we conducted a general linear model that involved a 2 between subjects (text format) \times 2 within subjects (confidence level: high vs. low) mixed design, which involved both performance on the true/false test and on the open-ended test. This analysis allowed for a test of the predicted interactions between text format and confidence level while taking into account any correlations among the dependent measures. First, there were significant main effects of text format, $F(1,64) = 4.63$, $p = .035$, $\eta_p^2 = .07$ and confidence level, $F(3, 192) = 32.39$, $p < .001$, $\eta_p^2 = .34$. These main effects indicate more correction of misconceptions after reading refutation texts than after reading standard texts, and more correction of low-confidence misconceptions than high-confidence misconceptions.

Fig. 2a illustrates the effects of confidence and text format on correction of misconceptions on the true/false test. Most importantly, as evident from inspection of Fig. 2a, the contrasts representing the interaction between Text Format and Confidence (high vs. low) for the true/false test was significant, $F(1,64) = 19.64$, $p < .001$, $\eta_p^2 = .24$.

Table 1

Percentage of overall performance, correction of misconceptions, magnitude of POPs, and selections for restudy.

	Refutation Text Group	Standard Text Group
Percentage Overall Test Performance		
Correct Responses True/False Statement Test	88.49 (12)	78.94 (14)
Misconceptions Open-Ended Test	13.04 (9)	19.85 (14)
Correct Responses Open-Ended Test	61.4 (15)	55.37 (15)
Correct Responses Detail Test	20.78 (39)	21.33 (40)
Percentage Corrected Misconceptions		
High-Confidence Misconceptions Correct on True/False Statement Test	87.84 (22)	63.60 (31)
Low-Confidence Misconceptions Correct on True/False Statement Test	91.93 (22)	91.89 (24)
High-Confidence Misconceptions Correct on Open-Ended Test	58.28 (32)	48.34 (29)
Low-Confidence Misconceptions Correct on Open-Ended Test	76.88 (37)	72.75 (39)
Monitoring and Regulation (POP Magnitude and Restudy Selections)		
POPs Misconceptions True/False Statement Test	53.88 (21)	54.71 (19)
POPs Correct Responses True/False Statement Test	60.74 (18)	59.77 (16)
POPs Misconceptions Open-Ended Test	56.54 (23)	54.53 (17)
POPs Correct Responses Open-Ended Test	62.12 (19)	59.89 (17)
Restudy Selections Misconceptions True/False Statement Test	32.74 (33)	32.41 (30)
Restudy Selections Correct Responses True/False Statement Test	17.70 (17)	21.84 (18)
Restudy Selections Misconception Open-Ended Test	23.64 (28)	29.40 (34)
Restudy Selections Correct Response Open-Ended Test	16.26 (19)	22.37 (20)

Note: Standard deviations of the Mean in parentheses.

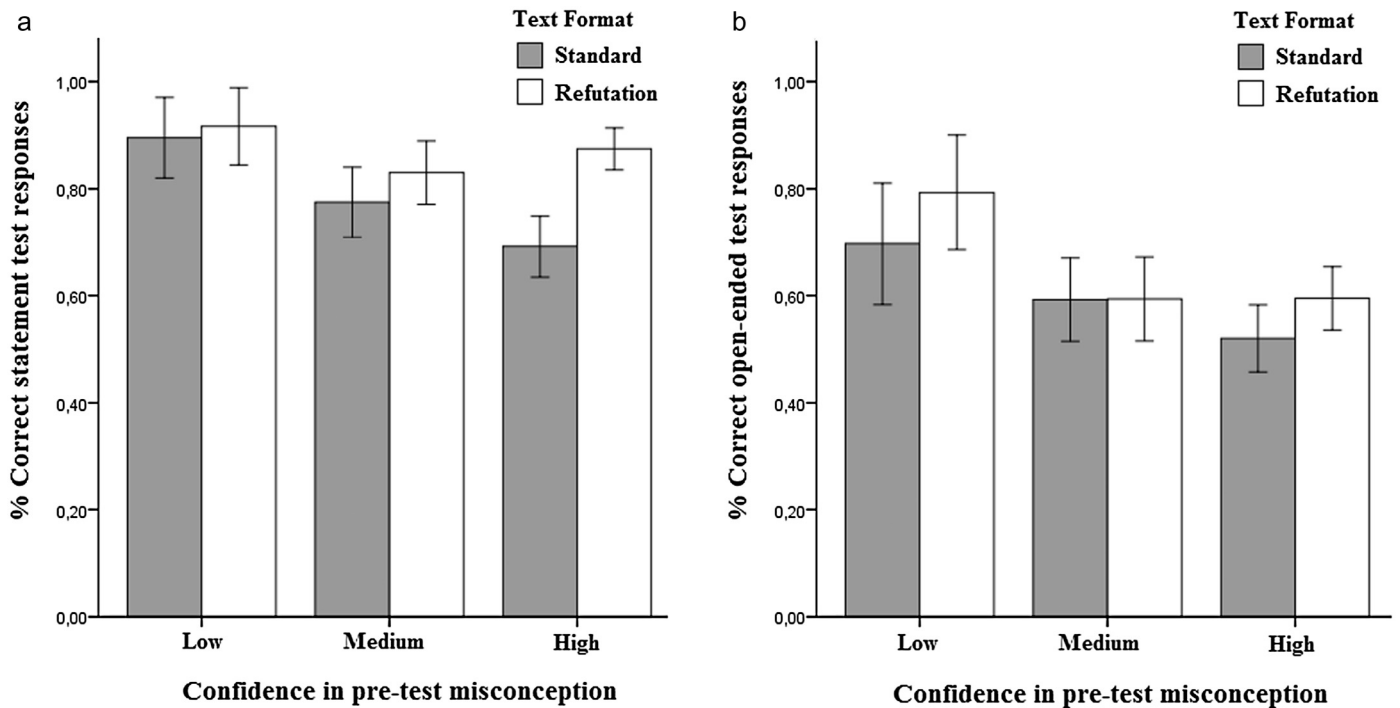


Fig. 2. Outdating and Updating of Misconceptions. Correct responses on the true/false statement test (a) and the open-ended test (b) as a function of confidence in the misconception on the pre-reading test. The error bars indicate the 95% confidence interval.

The univariate tests to follow up the interaction showed that reading refutation texts had a beneficial effect on correction of high-confidence misconceptions on the true/false test, $F(1, 64) = 19.34$, $p < .001$, $\eta_p^2 = .23$, whereas text format did not influence the correction of low-confidence misconceptions, $F(1, 64) = .044$, $p = .835$. Fig. 2b shows the effects of confidence and text format on correction of misconceptions on the open-ended test. Interestingly, as shown in Fig. 2b, text format differentially affected updating of high-confidence misconceptions on the open-ended test. In contrast to the true/false test, for the open-ended test (Fig. 2b), the focal Text Format by Confidence interaction was not significant, $F(1, 64) = 0.68$, $p = .68$. Thus, reading refutation texts did not lead to more updating of high-confidence misconceptions (as measured by the open-ended test) than reading standard texts.

In sum, these findings partially confirm the RCH-hypothesis and indicate that the benefit of reading refutation texts on correction of high-confidence misconceptions is more apparent for outdating of misconceptions than for updating misconceptions with correct information. Finally, to evaluate Hypothesis 2, we examined the influence of text format on performance on the detail test. As predicted, reading those texts did not lead to extra processing of details from the texts, $t(112) = .203$, $p = .84$.

3.2. POPs and restudy selections

In addition to evaluating the RCH hypothesis, we aimed to address whether reading refutation text fosters metacognition, by improving monitoring accuracy (Question 1) and study selections (Question 2) for misconceptions that are not yet corrected through study. Monitoring accuracy was measured by relating participants' POPs to their test performance. Participants' POPs ranged from 0% (indicating the learners are sure they will not be able to provide a correct test response) to 100% (confident to provide the correct test response). Table 1 shows the mean POPs for misconceptions and correct responses on the true/false test and the open-ended test. Note that, when monitoring for incorrect test responses is accurate, POPs should

be low in magnitude because learners will not receive credit for their incorrect test responses. If reading refutation texts improves monitoring accuracy for misconceptions that are not yet corrected, POPs should be more accurate (that is, less overconfident) after reading refutation texts than after reading standard texts. As shown in Table 1, the magnitude of the POPs for misconceptions on the true/false and the open-ended test indicates that learners were overconfident. Further, there were no large differences between POP magnitudes for the group reading refutation texts and the group reading standard texts. Furthermore, if learners' study selections are more adaptive after reading refutation texts, they should more often select the texts for which their misconceptions are not yet corrected. From Table 1, it appears that both the group reading refutation texts and the group reading standard texts did not select most of the texts for which they still held misconceptions for further study, indicating maladaptive self-regulation.

As per analyses of test performance above, we used a 2 (text format) \times 2 (metacognition measure: accuracy of POPs; restudy selection) GLM for repeated measures that included outcomes for the true/false test and for the open-ended test. Most importantly, the main effect of text format was not significant, $F(1, 79) = 0.31$, $p = .58$. Thus, even though reading refutation texts supported correction of misconceptions more than reading standard texts, this text format did not lead to improved monitoring accuracy or to more adaptive restudy selections.

3.3. Additional analyses

For completeness, additional statistics about reading times and correlations between the measured variables are presented. Mean reading times per text were 36.1 ($SD = 19.5$) seconds for standard texts, and 38.8 ($SD = 19.0$) seconds for refutation texts, $t(112) = 1.150$, $p = .253$. In addition, Table 2 includes the correlations (separately for each group) between pre-test misconceptions, confidence in misconceptions, reading times, POP magnitudes, restudy selections, and performance on the true/false test, open-ended test, and detail test.

Table 2

Correlations between the measured variables broken down by group.

Group	Group Reading Refutation Texts							
	Misconceptions at pre-test	Confidence in Misconception	Text reading time	Magnitude of POPs	Restudy Selections	True/False Test Performance	Open-Ended Test Performance	Detail Test Performance
Reading Standard Texts	Misconceptions at pre-test	-.060	.132	-.265*	.174	-.003	-.055	.104
	Confidence in Misconceptions	-.024	-.178	.458**	-.329*	-.121	.081	.034
	Text reading time	-.050	-.124	-.113	.197	.297*	.314*	.362**
	Magnitude of POPs	-.177	.361**	-.057	-.241	.198	.219	.209
	Restudy Selections	.172	.063	.162	-.060	.105	.035	-.101
	True/False Test Performance	-.251	.135	.378**	.344**	-.239	-.678**	.470**
	Open-Ended Test Performance	-.324*	.230	.279*	.317	-.101	.767**	-.613**
	Detail Test Performance	-.325*	-.036	.288*	.243*	-.014	.444**	.583**

Note: Correlations between the number of misconceptions at the pre-test, confidence ratings for misconceptions, reading times, magnitude of the POPs, percentage of restudy selections, and performance on the three test types (true/false statement test, open-ended test, detail test).

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

4. Discussion

With this experimental study, we investigated whether reading refutation texts, containing explicit feedback about incorrectness of a misconception, has beneficial effects on adolescents' correction (outdating and updating) of high-confidence science misconceptions on a delayed test. In particular, we evaluated the Refutation-Causes-Hypercorrection (RCH) hypothesis, which predicts that reading refutation texts will lead to enhanced correction of high-confidence misconceptions in comparison to reading standard texts, and that there will be less of an effect of text format on correction of low-confidence misconceptions (Hypothesis 1). Moreover, we hypothesized that reading refutation texts would specifically focus learners on correction of misconceptions, and that text format would not affect learning of details (Hypothesis 2). An additional aim was to investigate whether reading refutation texts fostered metacognition by improving monitoring accuracy (Question 1) and regulation of study (Question 2). We discuss the findings regarding each of these hypotheses and questions in turn.

A key question is whether hypercorrection will occur when incorrect knowledge is refuted in a text. In line with hypercorrection research (Butterfield & Metcalfe, 2001, 2006; Metcalfe et al., 2012; Metcalfe & Finn, 2011), we expected that the explicit feedback in the refutation texts would surprise learners and capture their attention, especially when these refutations contradict high-confidence misconceptions (given that such refutations would presumably be the most surprising).

The findings provided mixed support for the RCH-hypothesis (Hypothesis 1). Reading refutation texts led to more correction of high-confidence errors than reading standard texts when learners were tested with the true/false test. In contrast, when learners were asked to respond to open-ended questions, reading refutation texts did not lead to hypercorrection of high-confidence misconceptions. As hypothesized, for the low-confidence misconceptions reading refutation texts did not support correction more than reading standard texts, neither on the true/false test nor on the open-ended test. Furthermore, reading refutation texts focused learners specifically on correction of misconceptions, and not on learning of details (confirming Hypothesis 2).

A relatively unique feature of the present research was that it included both true/false and open-ended tests, which allowed us to explore both outdating and updating. Importantly, correcting a misconception involves both outdating of the misconception (which happens when a misconception becomes less accessible in memory because it has lost activation in the memory representation; Kendeou et al., 2013; Zwaan & Radvansky, 1998), and updating of the memory representation by the incorporation of correct information (encoding

and accommodating novel information in memory, Kendeou et al., 2013; Rapp & Kendeou, 2007). Comparing the effect of text format on responses to true/false statements and responses on the open-ended test allowed us to investigate both outdating and updating of misconceptions. When responding to the statements, learners were merely asked to indicate whether a stated fact was true or false, whereas when responding to the open-ended questions, learners were required to explain the correct information. Critically, reading refutation texts did not support correction of misconceptions on the true/false test and the open-ended test in the same manner. This indicates that correction is more apparent for outdating of misconceptions than for updating.

Interestingly, the correlation between outdating and updating misconceptions indicates a strong relation (as shown in Table 2, a strong correlation occurred between scores on the true/false test and the open-ended test), whereas (as just discussed) the effect of reading refutation texts had a stronger effect on outdating the high-confidence misconceptions on the true/false test than on updating the high-confidence misconceptions with correct information on the open-ended test. This might imply that outdating the misconception was more successful than updating the memory representation. Indeed, a follow-up analysis indicates that even when learners provided correct true/false test responses for initially incorrect misconceptions, they often failed to show that they accommodated the novel information into their prior knowledge when responding to open-ended questions. Both high-confidence misconceptions and low confidence-misconceptions were more often corrected on the true/false test than on the open-ended test (for high-confidence misconception $t(111) = 8.98, p < .001, r^2 = .42$; for low-confidence misconceptions $t(67) = 4.28, p < .001, r^2 = .21$). These findings suggest that learners processed the refutations in the text in a shallow fashion (Kendeou & van den Broek, 2005) and focused more on outdating than on updating. Future research should further investigate how to help learners to update. Perhaps providing learners with a more detailed explanation about the correct information might lead to deeper processing and more successful updating (Kendeou, Walsh, Smith, & O'Brien, 2014).

This study was the first to investigate hypercorrection in adolescent learners when reading complex science texts, in contrast to research on hypercorrection using short general information questions (Metcalfe & Finn, 2012). Even though the present study shows that feedback provided through refutations supported correction of high-confidence misconceptions, we did not replicate the classical hypercorrection effect; namely, in the present research, the high-confidence errors were less likely to be corrected than were low-confidence errors (cf. Ecker et al., 2010; Pintrich et al., 1993). In contrast, in prior hypercorrection studies (Butterfield & Metcalfe,

2001, 2006; Fazio & Marsh, 2009; Metcalfe et al., 2012; Metcalfe & Finn, 2011), high-confidence errors are corrected more often than are low confidence errors. The difference between the current findings and the prior hypercorrection research could possibly be explained by differences in the feedback format. In the present experiment, learners received delayed feedback about the correctness of pre-reading test responses through reading of the expository science texts, whereas in prior hypercorrection studies, learners received direct feedback in the form of short sentences containing the correct answer. Further, in prior hypercorrection studies, learners only received detailed feedback about the correct response when they made an error, in contrast to the present study where readers read texts about all topics, whether they held misconceptions about these topics or not. Therefore, in the current study, learners' attention might not have been captured as much by the feedback as it was in the original hypercorrection studies. Future research should shed more light on the effects of different kinds (and amount) of feedback on hypercorrection of high-confidence errors.

The majority of research on refutation texts focused on improvement of performance and it was not known whether or how the reading of refutation texts affects metacognition. Thus, we also addressed whether reading refutation texts would improve monitoring accuracy (Question 1) and restudy selections (Question 2). Our findings show that reading refutation texts did not have any beneficial effect on monitoring accuracy: Participants were highly overconfident when predicting performance for misconceptions that they would provide on the test. Learners base their monitoring of learning on retrieval of accessible information from long-term memory (Rhodes & Tauber, 2011). Possibly, because learners could retrieve misconceptions from memory, the accessibility of information was not predictive of correct test performance (Koriat, 1998). This might have led to maladaptive regulation; when selecting texts for restudy, they prematurely discarded most of the texts about which they still held misconceptions.

Results bring novel insight into the way misconceptions are dispelled when reading science texts. Overall, our findings confirm that reading refutation texts is a powerful method to improve correction of misconceptions. Importantly, individual differences in confidence in misconceptions affect correction, and the findings indicate that when learners are confident in the correctness of their misconceptions, they will need support to correct them. Furthermore, from this research it can be concluded that, when aiming to investigate and foster learners' correction of science misconceptions, it is important to take into account that results depend on the assessment method. Typically, conceptual change has been measured with multiple-choice tests; and even though these tests are suitable to measure outdated, only using multiple-choice measures might not provide detailed insight into learners' actual level of conceptual change taking place. To accurately investigate learners' correction of misconceptions, assessment should also include whether students have accommodated the novel information. We made use of open-ended questions to interpret and evaluate students' updating processes, and from our findings, it appears that updating is difficult to achieve even if outdated occurs. Because the updating task is also the most important task when applying knowledge in real life situations, we endorse that research on refutation texts should extend its assessment of correction from only investigating the easier task of outdated to also the additional process of updating.

Using multiple science texts provided insights into correction of misconceptions across texts, making the findings relevant for the educational setting where students are typically required to study various texts. Further, testing students after a delay improves the applicability of these findings in the educational context, where a focus is on comprehension and retention of knowledge over time, rather than on immediate testing.

4.1. Limitations and future directions

The results suggest that reading refutation texts surprised learners and drew their attention when they held misconceptions with high confidence. Note that even though we expected that the feedback about the misconceptions in the refutation text draws learners' attention because the inconsistency surprises them (based on research by Broughton et al., 2010; Butterfield & Mangels, 2003; Butterfield & Metcalfe, 2006; Cordova et al., 2014; Fazio & Marsh, 2009; Metcalfe et al., 2012), a limitation of the present study is that we did not measure learners' surprise. Thus, our account remains speculative. Future research could investigate the level of learners' surprise when reading refutation texts about misconceptions, such as by asking them to rate their surprise after reading each text (cf. Cordova et al., 2014) or perhaps measuring their pupil dilation using eye tracking technology to measure their arousal while reading the refutation.

Even though reading times did not differ for the group reading refutation texts and the group reading standard texts, the refutation texts were somewhat longer than the standard texts. A general limitation of the present study is that we do not have insight in the differences in reading processes between the two groups. For future research, it would be interesting to investigate the on-line reading processes (cf. Ariasi & Mason, 2011, 2014). Furthermore, we only measured off-line monitoring in which learners predicted their performance after studying the texts. This type of monitoring is important in the process of self-regulated learning, because it gives learners input for their restudy decisions (Metcalfe & Finn, 2008; Thiede & Dunlosky, 1999). A limitation of only using a measure of off-line monitoring is that it does not provide fine-grained measures about the metacognitive processes occurring during reading. For future research, it would be interesting to combine measures of off-line monitoring with measures of on-line monitoring, such as using think-aloud methods or investigating sentence-by-sentence reading times (cf. Kendeou & van den Broek, 2005).

Even though we consider using multiple texts a strength of this research, note that internal consistency between the texts was low. This low consistency was not altogether unexpected because the text topics were unrelated to each other, so variance in performance on the questions across the 16 texts need not be high (e.g., an individual may understand one topic and perform relatively well yet not understand another topic and perform poorly). Such variability across texts within an individual is desirable for estimating within-participant judgment accuracy, but it also is a potential limit of this study. Accordingly, to test generalizability of this research, we recommend replication by other researchers using a new set of texts. Additionally, in the present study, all learners followed education of the highest level (out of four levels) of Dutch secondary education. For future research on correction of misconceptions, it would be interesting to sample participants with a wider range of cognitive abilities.

4.2. Conclusions

This study adds to the growing literature on the hypercorrection effect and the effect of refutations on correcting misconceptions (e.g., Braasch et al., 2013; Mason et al., 2008; Metcalfe et al., 2012). Findings show that misconceptions about science are often not corrected when learners hold these with high confidence, but that reading refutation text does lead to more corrections of the high-confidence misconceptions as compared to reading standard texts. When learners experience cognitive dissonance, because misconceptions are co-activated with the correct information, this clearly helps them to outdated the misconceptions that are held with high confidence. However, providing learners with refutation texts does not seem sufficient for the updating process; that is, co-activation was not sufficient to support

the learners to accommodate the novel, contrasting information in their pre-existing knowledge base. Further, reading refutation texts was not sufficient to improve the self-regulated learning process for the misconceptions that were not corrected through extra study. Thus, in terms of Nelson and Narens (1990), these findings seem to imply that reading refutation texts has more effect on the cognitive level than on the metacognitive level.

Importantly, these findings have implications for classroom practice, by showing that explicit refutation of the misconception supports correction of misconceptions that are strongly held. Such corrections were evident one week after the students' misconceptions were refuted. Refutation texts are a powerful tool to help students outdate misconceptions, but additional interventions may be needed to improve learners' accommodation of the contrasting, correct information.

Acknowledgment

This research was supported by the Netherlands Organisation for Scientific Research (NWO), grant 411.07.151. The authors would like to thank Gerrit Drost for programming of the experimental tasks and support with data analyses.

Appendix A: Misconceptions Addressed in this Study

Misconception	Correct Information
1. The largest part of the Sahara consists of sand	1. The largest part of the Sahara consists of rocks
2. Most male animals are stronger than the females	2. Most female animals are stronger than the males
3. Columbus discovered North America	3. The Vikings made settlements long before Columbus discovered America
4. Bees always die after stinging	4. Bees can stay alive after stinging
5. Bats are blind	5. Bats can see (and catch prey by using vision)
6. Seasons are caused by the earth being closer to the sun in summer than in winter	6. Seasons are caused by the tilted position of the earth (and the sun is even further away from the earth in summer)
7. Dragonflies can sting	7. Dragonflies can't sting
8. Spanish peppers originally come from Spain	8. Spanish peppers originally come from Mexico
9. A camel stores water in its humps	9. A camel stores fat in its humps
10. Objects float in water because they are lighter than water	10. Objects float in water because they have buoyancy and can move water
11. You get a cold from exposure to cold temperatures	11. You get a cold from a virus
12. Eskimos have 20 different words for snow	12. Eskimos don't have more words for snow than we do
13. A falling star is a star	13. A falling star is not a star
14. The brightest stars are the largest stars	14. The brightest stars are closer to the Earth
15. It takes years to digest swallowed chewing gum	15. Chewing gum does not remain in the body, so there is no need for digestion
16. When in a heavy thunderstorm, it is safest to lay down flat on the ground	16. When in a heavy thunderstorm, it is safest to get into a squatted position

Appendix B: Examples of Standard Texts and Refutation Texts

Refutation Text Star Brightness

In ancient times, the scientist Hipparchus defined the brightness of stars with numbers. Nowadays, scientists use the term 'Magnitude' to define the brightness of stars.

Not all stars are at the same distance to the Earth, some stars are closer to us, and some are further away. A lot of people think that the brightest stars are the largest stars, but this is not necessarily true. The brightest stars are closer to Earth than stars that are less bright.

The star that's nearest to the Earth is the Sun. The Sun is an average size star. It is 190 times bigger than Earth, and the distance from Earth to the Sun is approximately 150 million kilometers.

Standard Text Star Brightness

In ancient times, the scientist Hipparchus defined the brightness of stars with numbers. Nowadays, scientists use the term 'Magnitude' to define the brightness of stars.

Not all stars are at the same distance to the Earth, some stars are closer to us, and some are further away. The brightest stars are closer to Earth than stars that are less bright.

The star that's nearest to the Earth is the Sun. The Sun is an average size star. It is 190 times bigger than Earth, and the distance from Earth to the Sun is approximately 150 million kilometers.

Refutation Text Stinging Bees

Bees have stingers that are strongly barbed. When they sting, the venom is secreted from the venom gland into the skin. Only female bees have stingers.

A lot of people think that bees always die after stinging. In most cases, this is not true. When bees sting, they most often sting insects or other animals that, just like insects, have an external skeleton made out of chitin. It protects these animals and gives them stability. Despite the barbs, bees can retract their stingers from these animals.

Only when stinging animals with a thick, flexible skin, bees most of the time get stuck and lose their stinger after stinging. This causes them to lose part of their rear body, leaving a big wound, which most of the time causes their death.

Standard Text Stinging Bees

Bees have stingers that are strongly barbed. When they sting, the venom is secreted from the venom gland into the skin. Only female bees have stingers.

When bees sting, they most often sting insects or other animals that, just like insects, have an external skeleton made out of chitin. It protects these animals and gives them stability. Despite the barbs, bees can retract their stingers from these animals.

Only when stinging animals with a thick, flexible skin, bees most of the time get stuck and lose their stinger after stinging. This causes them to lose part of their rear body, leaving a big wound, which most of the time causes their death.

Appendix C: Examples of Post-Test Questions

Open-Ended Questions about Misconceptions

- Why are some stars brighter than other stars? (Correct response: The brighter stars are closer to the Earth.)
- What happens when a bee stings another insect? (Correct response: The bee can retract its stinger and will survive.)

True/False Statement Questions

- The brighter stars are closer to the Earth than stars that are less bright. (True)
- A bee always dies after stinging. (False)

Detail Test Questions

- Which term is used by scientists to define the brightness of stars? (Correct response: Magnitude.)
- Which bees have stingers? (Correct response: Female bees.)

References

- Alvermann, D. E. (2002). Effective literacy instruction for adolescents. *Journal of Literacy Research*, 34(2), 189–208. doi:10.1207/s15548430jlr3402_4.
- Alvermann, D. E., & Hynd, C. R. (1989). Effects of prior knowledge activation modes and text structure on nonscience majors' comprehension of physics. *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, 581–601. doi:10.1007/s11251-010-9142-5.
- Ariasi, N., & Mason, L. (2014). From covert processes to overt outcomes of refutation text reading: The interplay of science text structure and working memory capacity through eye fixations. *International Journal of Science and Mathematics Education*, 12, 493–523. doi:10.1007/s10763-013-9494-9.
- 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(3), 561–578. doi:10.1037/a0032627.
- Broughton, S. H., Sinatra, G. M., & Reynolds, R. E. (2010). The nature of the refutation text effect: An investigation of attention allocation. *Journal of Educational Research*, 103, 407–423. doi:10.1080/00220670903383101.
- Butler, A. C., Fazio, L. K., & Marsh, E. J. (2011). The hypercorrection effect persists over a week, but high-confidence errors return. *Psychonomic Bulletin & Review*, 18(6), 1238–1244. doi:10.3758/s13423-011-0173-y.
- Butterfield, B., & Mangels, J. A. (2003). Neural correlates of error detection and correction in a semantic retrieval task. *Brain Research. Cognitive Brain Research*, 17(3), 793–817. doi:10.1016/s0926-6410(03)00203-9.
- Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 27(6), 1491–1494. doi:10.1037/0278-7393.27.6.1491.
- Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, 1(1), 69–84. doi:10.1007/s11409-006-6894-z.
- Cordova, J. R., Sinatra, G. M., Jones, S. H., Taasobshirazi, G., & Lombardi, D. (2014). Confidence in prior knowledge, self-efficacy, interest and prior knowledge: Influences on conceptual change. *Contemporary Educational Psychology*, 39, 164–174. doi:10.1016/j.cedpsych.2014.03.006.
- Diakidoy, I. A. N., Kendeou, P., & Ioannides, C. (2003). Reading about energy: The effects of text structure in science learning and conceptual change. *Contemporary Educational Psychology*, 28(3), 335–356. doi:10.1016/s0361-476x(02)00039-5.
- Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, 33(2–3), 109–128. doi:10.1207/s15326985sep3302&3_5.
- Dunlosky, J., Hartwig, M. K., Rawson, K. A., & Lipko, A. R. (2011). Improving college students' evaluation of text learning using idea-unit standards. *Quarterly Journal of Experimental Psychology* (2006), 64(3), 467–484. doi:10.1080/17470218.2010.502239.
- Dunlosky, J., & Thiede, K. W. (2013). Four cornerstones of calibration research: Why understanding students' judgments can improve their achievement. *Learning and Instruction*, 24, 58–61. doi:10.1016/j.learninstruc.2012.05.002.
- Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review*, 18(3), 570–578. doi:10.3758/s13423-011-0065-1.
- Ecker, U. K. H., Lewandowsky, S., & Tang, D. T. W. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, 38(8), 1087–1100. doi:10.3758/mc.38.8.1087.
- Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, 16(1), 88–92. doi:10.3758/pbr.16.1.88.
- Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science—A comparative metaanalysis of instructional interventions from reading education and science-education. *Reading Research Quarterly*, 28(2), 117–155. doi:10.1002/acp.1418.
- Hammer, D. (1996). More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research. *American Journal of Physics*, 64(10), 1316–1325. doi:10.1119/1.18376.
- Kendeou, P., & O'Brien, E. J. (2014). The Knowledge Revision Components (KReC) framework: Processes and mechanisms. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 353–377). Cambridge, MA: MIT Press.
- Kendeou, P., Smith, E. R., & O'Brien, E. J. (2013). Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 39(3), 854–865. doi:10.1037/a0029468.
- Kendeou, P., & van den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology*, 97(2), 235–245. doi:10.1037/0022-0663.97.2.235.
- Kendeou, P., Walsh, E. R., Smith, E. R., & O'Brien, E. J. (2014). Knowledge revision processes in refutation texts. *Discourse Processes*, 51, 1–24. doi:10.1080/0163853X.2014.913961.
- Koriat, A. (1998). Illusions of knowing: The link between knowledge and metaknowledge. In V. Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Cognitive and social dimensions* (pp. 16–34). Thousand Oaks, CA: SAGE.
- Mason, L., Gava, M., & Boldrin, A. (2008). On warm conceptual change: The interplay of text, epistemological beliefs, and topic interest. *Journal of Educational Psychology*, 100(2), 291–309. doi:10.1037/0022-0663.100.2.291.
- McCloskey, M. (1983). Naive theories of motion. In D. G. A. L. Stevens (Ed.), *Mental models* (pp. 299–324). Hillsdale, NJ: Erlbaum.
- Metcalfe, J., Butterfield, B., Habeck, C., & Stern, Y. (2012). Neural correlates of people's hypercorrection of their false beliefs. *Journal of Cognitive Neuroscience*, 24(7), 1571–1583.
- Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, 15, 174–179. doi:10.3758/PBR.15.1.174.
- Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high-confidence errors: Did they know it all along? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 37(2), 437–448. doi:10.1037/a0021962.
- Metcalfe, J., & Finn, B. (2012). Hypercorrection of high confidence errors in children. *Learning and Instruction*, 22(4), 253–261. doi:10.1016/j.learninstruc.2011.10.004.
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). New York: Academic Press.
- Otero, J., Leon, J. A., & Graesser, A. C. (2002). Introduction to the psychology of science text comprehension. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 1–15). Mahwah, NJ: Erlbaum.
- Piaget, J. (1964). Part 1, Cognitive development in children: Development and learning. *Journal of Research in Science Teaching*, 2, 176–186.
- Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, 63, 167–199. doi:10.3102/0034654306300216.
- Rapp, D. N., & Kendeou, P. (2007). Revising what readers know: Updating text representations during narrative comprehension. *Memory & Cognition*, 35(8), 2019–2032. doi:10.3758/bf03192934.
- Rawson, K. A., & Dunlosky, J. (2013). Relearning attenuates the benefits and costs of spacing. *Journal of Experimental Psychology. General*, 142(4), 1113–1129. doi:10.1037/a0030498.
- Rhodes, M. G., & Tauber, S. K. (2011). Monitoring memory errors: The influence of the veracity of retrieved information on the accuracy of judgements of learning. *Memory (Hove, England)*, 19(8), 853–870. doi:10.1080/09658211.2011.613841.
- Sanchez, E., & Garcia-Rodicio, H. (2013). Using online measures to determine how learners process instructional explanations. *Learning and Instruction*, 26, 1–11. doi:10.1016/j.learninstruc.2012.12.003.
- Schneps, M. (1989). *A private universe*. San Francisco: Astronomical Society of the Pacific. <http://www.youtube.com/watch?v=TrXaQu_qGeo> Accessed 04.05.15.
- Sinatra, G. M., & Broughton, S. H. (2011). Bridging reading comprehension and conceptual change in science education: The promise of refutation text. *Reading Research Quarterly*, 46(4), 374–393. doi:10.1002/rrq.005.
- Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 26, 204–221.
- Thiede, K. W., Anderson, M. C. M., & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, 95(1), 66–73. doi:10.1037/0022-0663.95.1.66.
- Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 25, 1024–1037. doi:10.1037/0278-7393.25.4.1024.
- Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. (2009). Metacognitive monitoring during and after reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 85–106). UK: Routledge.
- 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. doi:10.1007/s10763-010-9203-x.
- 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. doi:10.1002/acp.1418.
- Van Loon, M. H., de Bruin, A. B. H., van Gog, T., & van Merriënboer, J. J. G. (2013). Activation of inaccurate prior knowledge affects primary-school students' metacognitive judgments and calibration. *Learning and Instruction*, 24, 15–25. doi:10.1016/j.learninstruc.2012.08.005.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162–185. doi:10.1037/0278-7393.24.5.1224.