Refuting Overconfidence: Refutation Texts Prevent Detrimental Effects of Misconceptions on Text Comprehension and Metacomprehension Accuracy in the Domain of Statistics

Anja Prinz (anja.prinz@ezw.uni-freiburg.de)
Department of Educational Science, University of Freiburg
Rempartstraße 11, 79098 Freiburg, Germany

Stefanie Golke (stefanie.golke@ezw.uni-freiburg.de)
Department of Educational Science, University of Freiburg
Rempartstraße 11, 79098 Freiburg, Germany

Jörg Wittwer (joerg.wittwer@ezw.uni-freiburg.de)
Department of Educational Science, University of Freiburg
Rempartstraße 11, 79098 Freiburg, Germany

Abstract

Refutation texts are beneficial for removing misconceptions and supporting comprehension in science. Whether these beneficial effects hold true in the domain of statistics is, however, an open question. Moreover, the role of refutation texts for the accuracy in judging one's own comprehension (metacomprehension accuracy) has received little attention. Therefore, we conducted an experiment in which students with varying levels of statistical misconceptions read either a standard text or a refutation text in statistics, judged their text comprehension, and completed a comprehension test. The results showed that when students read the standard text, having more misconceptions resulted in poorer text comprehension and more inaccurate metacomprehension as indicated by overconfident predictions. In contrast, when students read the refutation text, the number of misconceptions was unrelated to text comprehension and metacomprehension accuracy. Apparently, refutation texts help students to pay attention to inaccuracies in their knowledge and, thereby, can promote self-regulated learning from texts.

Keywords: metacomprehension accuracy; misconceptions; procedural and conceptual understanding; text comprehension

In higher education, statistics has become a central part in many fields of study to enable students to deal with quantitative information (Ben-Zvi & Garfield, 2004). At the same time, in higher education, students are increasingly expected to engage in self-regulated learning (Cassidy, 2011). For example, in statistics, students often need to advance their knowledge by reading statistics textbooks. However, such learning can be challenging, especially when students have to understand complex statistical concepts, such as covariance, about which they frequently have false ideas in the form of misconceptions. Such statistical misconceptions differ in fundamental ways from the normatively correct conceptions and, thus, can strongly hamper the comprehension and application of statistics (Liu, 2010).

The Role of Misconceptions for Text Comprehension and Metacomprehension

Text comprehension is a process by which learners actively construct a mental representation of the information provided in a text (Kintsch, 1998). As usually not all possible relations are explicitly stated in a text, the construction of a mental representation requires learners to use their prior knowledge to infer within-text and knowledge relations (McNamara & Magliano, 2009).

However, when learners possess inaccurate prior knowledge in the form of misconceptions, comprehension can be hampered because misconceptions can trigger false inferences. For example, Kendeou and van den Broek (2005) examined the online processes taking place when learners with misconceptions read texts. Their findings showed that learners with misconceptions used and integrated their prior knowledge with textual information as did learners without misconceptions. Yet, the content of their inferences was contaminated by their misconceptions. This, in turn, resulted in an inappropriate mental text representation after reading. In line with these findings, research in reading and science education shows that misconceptions often hinder memory and comprehension of text (see, e.g., Guzzetti et al., 1993).

When learning from reading texts, it is also important that learners accurately monitor and judge their own comprehension, which is known as metacomprehension accuracy (Dunlosky & Lipko, 2007). Accurate metacomprehension affects the extent to which learners effectively self-regulate their learning. For example, only a learner who accurately monitors that a text is not yet sufficiently understood to perform well on a comprehension test might decide to further study the material (e.g., Thiede, Anderson, & Therriault, 2003). Research, however, indicates that learners are often overconfident when monitoring their text comprehension. This is particularly true for learners' predictions, that is, their judgments of comprehension after they have read a text but before they have taken a comprehension test (e.g., Maki et al., 2005).

Recently, Prinz, Golke, and Wittwer (2017) found that misconceptions produced overconfident predictions. More precisely, their results showed that statistical misconceptions not only impeded the comprehension of a statistics text but also led to inaccurate self-assessments of text comprehension as indicated by overconfident predictions. Apparently, when learners have misconceptions, they are likely to construct a flawed mental text representation. At the same time, when self-assessing their text comprehension, learners tend to focus on the amount of textual information they can retrieve from memory while neglecting whether this information is correct. Consequently, learners with misconceptions are likely to overestimate their actual comprehension (see also Dunlosky, Rawson, & Middleton, 2005).

In sum, research indicates that learning from standard texts mainly elicits superficial understanding and monitoring when learners hold misconceptions. To overcome these difficulties, it is important that learners become aware of their misconceptions and revise their understanding, a process called conceptual change.

The Role of Refutations for Text Comprehension and Metacomprehension

Conceptual change occurs when learners modify their existing prior knowledge to include new information (e.g., Chi, 2008). This requires that the existing prior knowledge is identified as inadequate and the new information is understandable, plausible, and useful (Posner et al., 1982). Usually, conceptual change is demanding because a strongly held misconception impedes the recognition of its inconsistency with the correct information provided in a text (Otero & Kintsch, 1992). A promising instructional approach to inducing conceptual change is the use of refutation texts (Guzzetti et al., 1993). Refutation text passages typically comprise three elements: First, a commonly held misconception is described. Second, a cue that explicitly states that the misconception is in fact inaccurate is presented. Third, the scientifically correct explanation that directly refutes the misconception is provided (Tippett, 2010).

Research has shown that refutation texts in science domains are indeed more beneficial for restructuring incorrect prior knowledge than standard expository texts (e.g., Ariasi & Mason, 2011; see also Guzzetti et al., 1993; Tippett, 2010). Studies that investigated the processes taking place when reading refutation texts found that refutation texts are effective because misconceptions and correct conceptions are presented in close proximity, thereby increasing the likelihood of simultaneous activation. Only after such coactivation, further conceptual change processes, like the experience of a cognitive conflict, the evaluation of one's current conceptions, and the establishment of coherence in one's knowledge, can take place (e.g., Ariasi & Mason, 2011; van den Broek & Kendeou, 2008). Whether refutation texts prove effective in promoting conceptual change and enhancing comprehension also in statistics is not yet clear. Learning in statistics typically involves the acquisition of both concepts and procedures (Ben-Zvi & Garfield, 2004). Therefore, it is also an open question whether refutation texts not only benefit the learning of statistical concepts but also the learning of statistical procedures.

With regard to metacomprehension, the role of refutation texts is largely under-researched. An exception is a study conducted by van Loon et al. (2015) that revealed no beneficial effects of refutation texts on monitoring accuracy because learners remained overconfident when predicting their comprehension. However, the texts about misconceptions used in the study were rather short and there was only one comprehension question per text. Therefore, predictions were related exclusively to the comprehension of information about a single misconception. Hence, it is unclear whether refutation texts would promote metacomprehension accuracy when judgments do not focus exclusively on misconceptions. In literature on conceptual change, it has often been theorized that refutation texts increase learners' metacognitive awareness of their own conceptions in relation to the scientific conceptions (e.g., Ariasi & Mason, 2011). Thus, it seems plausible to assume that refutation texts can support learners in reflecting about their misconceptions, thereby increasing metacomprehension accuracy.

The Present Study

We investigated to what extent a refutation text compensates for the detrimental impact of misconceptions on text comprehension and metacomprehension accuracy in the domain of statistics. More specifically, we focused on the topic of covariance and examined comprehension and metacomprehension accuracy with respect to both conceptual and procedural aspects of covariance.

The first research question addressed whether the type of text would moderate the effect of misconceptions on text comprehension. We expected that when reading a standard text, more misconceptions would lead to poorer conceptual and procedural text comprehension, whereas this relationship would not be apparent when reading a refutation text.

The second research question concerned whether the type of text would moderate the effect of misconceptions on metacomprehension accuracy. We hypothesized that when reading a standard text, more misconceptions would lead to greater overestimation of conceptual and procedural text comprehension, whereas this detrimental effect would not be apparent when reading a refutation text.

Method

Participants and Design

A total of N = 53 university students (M = 25.04 years, SD = 2.42, 59% female) participated in this study. The study had two independent variables. The first independent variable was categorical and referred to the type of text: Participants read either a standard text or a refutation text about the statistical concept of covariance. The second independent variable was metric and referred to the number of misconceptions about covariance. Dependent variables were text comprehension and metacomprehension accuracy referring to conceptual and procedural aspects of covariance.

Material

The statistics text about covariance was adapted from a statistics textbook written by Bortz and Schuster (2010). This text existed in two versions: a standard text version and a refutation text version. Both versions addressed conceptual aspects of covariance such as its different directions, explained procedural aspects of covariance such as how it is calculated, provided the formula for computing covariance, and contained three graphs to illustrate positive, negative, and no covariance. In addition, the refutation text contained information challenging four common misconceptions about covariance (i.e., covariance implies causality, covariance is a standardized statistic, covariance is related to the slope of the fit line, and zero covariance proves the absence of any association; see, e.g., Prinz et al., 2017). More precisely, for each misconception, the three typical elements of a refutation text passage were provided: First, the misconception was described. Second, the incorrectness of the misconception was explicitly stated. Third, the scientifically correct explanation was given (Tippett, 2010). In contrast, the standard text only provided the scientifically correct explanation for each misconception. Without the graphs and the formula, the standard text included 515 words and the refutation text included 638 words. We did not equate the length of the two text versions to keep the manipulation unconfounded with other variations (e.g., the inclusion of additional or repetitive information in the standard text; cf., e.g., Diakidoy, Mouskounti, & Ioannides, 2011).

Measures

Misconceptions Misconceptions about covariance were assessed by 15 questions, with each question addressing one particular misconception. We collected these misconceptions on the basis of a comprehensive literature review (Prinz et al., 2017). For example, one question referred to the misconception that covariance does not depend on measurement units but represents a standardized statistic:

> In a study, sports scientists from a university determined the covariance between the height and the time for a 100-m dash of 20 sprinters. In his calculation, sports scientist A quantified time in seconds. When his colleague, sports scientist B, checks again, he quantifies time in milliseconds. Which of the following statements about the covariances calculated by the two sports scientists is correct?

- ☐ The two sports scientists will receive the same covariance because it does not matter if they use different measurement units (misconception).
- ☐ Both calculations will yield no covariance because one cannot calculate covariance from time data (wrong).
- ☐ No statement about the two covariances can be made because it is unknown if the variables time and height are linear (wrong).
- ☐ Sports scientist B will obtain a higher covariance than sports scientist A because milliseconds yield bigger numbers than seconds (correct).

All questions had a single-choice format with four response options. One option represented the correct answer, one option represented the particular misconception, and the two remaining options represented incorrect answers but not a particular misconception. The number of misconceptions was determined by counting how many times participants selected the response option that represented a misconception. Thus, they could record a maximum number of 15 misconceptions.

Text Comprehension Text comprehension referred to both conceptual and procedural comprehension of covariance. Conceptual comprehension was assessed by eight inference questions that had a single-choice format with four response options. Of the eight questions, four questions addressed misconceptions about covariance as already described. These were the four misconceptions that were targeted in the text. For these questions, one response option represented the correct answer, one response option represented the particular misconception, and the two remaining response options represented incorrect answers but not a particular misconception. Another four questions addressed further conceptual attributes of covariance but not specifically misconceptions. For these questions, one response option represented the correct answer and three response options represented incorrect answers but not a particular misconception. The participants received 1 point for each correct answer. Thus, they could achieve a maximum number of 8 points in the conceptual comprehension test.

Procedural comprehension was assessed by four open-ended questions that required the participants to perform calculations regarding covariance. They received 1 point for each correct answer. Thus, they could achieve a maximum number of 4 points in the procedural comprehension test. Interrater agreement on the procedural comprehension questions was high, Cohen's $\kappa = .98$, 95% CI [0.95, 1.00].

To facilitate the interpretation of participants' performance on the conceptual and procedural comprehension questions, we converted the number of conceptual and procedural comprehension questions correct into percent correct.

Metacomprehension Accuracy Before completing the comprehension questions, participants predicted the number of questions they would presumably answer correctly. They made their predictions for the conceptual and procedural questions separately. Metacomprehension accuracy was calculated by taking the signed difference between participants' judged number of questions correct (converted into percent correct) and their actual number of questions correct (converted into percent correct). Hence, a positive value indicated overconfidence because participants would assume to answer more comprehension questions correctly than they actually did. For example, a value of +.10 means that participants assumed to provide 10% more correct answers to the questions than they actually did. In contrast, a negative value indicated underconfidence and a value of zero indicated a perfectly accurate judgment.

Procedure

In the experiment, first, the participants completed the misconceptions test about covariance. Second, they accomplished a reading skills test serving as a filler task to remove the contents of the misconceptions test from working memory. Third, the participants read the statistics text about covariance. They were informed that their conceptual and procedural comprehension of the text would be tested after reading. Fourth, the participants predicted their conceptual and procedural text comprehension. To do so, they were informed about what kind of knowledge the two types of comprehension questions would require. Fifth, they answered the conceptual and procedural comprehension questions.

Results

To statistically test our hypotheses, we performed multiple regressions. We centered all predictor variables to maintain meaningful estimates of the main effects. In case of a statistically significant interaction effect, we computed simple slopes analyses following the approach suggested by Richter (2007) to investigate the pattern of the interaction. According to this approach, the categorical predictor *text type* was dummy coded and entered in two complementary regression models to estimate the regression parameters. As before, the metric predictor *number of misconceptions* was entered in the regression models in centered form. When testing directional hypotheses, we used one-tailed tests. Table 1 displays descriptive statistics for misconceptions and the dependent variables as a function of text type.

Table 1: Descriptive statistics.

	Refutation text (n = 27)		Standard text (n = 26)		Total sample (N = 53)	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*
NoM	4.74	1.66	4.15	1.38	4.45	1.54
CC	.62	.19	.65	.20	.63	.19
PC	.59	.32	.65	.34	.62	.33
Accuracy CC	.07	.19	.07	.18	.07	.19
Accuracy PC	-.10	.37	-.14	.31	-.12	.34

Note. NoM = number of misconceptions; CC = conceptual comprehension; PC = procedural comprehension.

The refutation text group and the standard text group did not significantly differ from each other with regard to the number of misconceptions, $t(51) = -1.40$, $p = .167$, $d = 0.39$.

Text Comprehension

As displayed in Table 2, the multiple regression with conceptual comprehension revealed a marginal significant main effect of misconceptions and a significant interaction effect between text type and misconceptions.

Table 2: Predictors of conceptual comprehension.

Predictor	*b*	*SE b*	*t*(49)	*p*	ΔR^2
Constant	0.62	0.03	24.64	<.001	
Text type	-0.01	0.05	-0.18	.857	.01
NoM	-0.03	0.02	-1.88	.066	.03
Text type x NoM	0.09	0.03	2.73	.005	.13

Note. $R^2 = .17$, $F(3, 49) = 3.26$, $p = .029$. NoM = number of misconceptions.

Simple slopes analyses showed that, in the standard text group, the regression coefficient *b* for number of misconceptions was -0.08 (SE = 0.03) and significantly different from zero, $t(52) = -3.01$, $p = .002$, $\Delta R^2 = .03$. This means that an increase of one misconception led to a decrease of 8% in conceptual text comprehension. In contrast, in the refutation text group, the regression coefficient for number of misconceptions was not significant, $b = 0.01$, $SE = 0.02$, $t(52) = 0.64$, $p = .527$, $\Delta R^2 = .03$. Thus, there was no significant effect of misconceptions.

As shown in Table 3, the multiple regression with procedural comprehension also revealed a significant interaction effect between text type and misconceptions.

Table 3: Predictors of procedural comprehension.

Predictor	*b*	*SE b*	*t*(49)	*p*	ΔR^2
Constant	0.61	0.05	13.38	<.001	
Text type	-0.04	0.09	-0.45	.654	.01
NoM	-0.03	0.03	-1.08	.287	.01
Text type x NoM	0.11	0.06	1.75	.043	.06

Note. $R^2 = .08$, $F(3, 49) = 1.38$, $p = .260$. NoM = number of misconceptions.

Simple slopes analyses showed that, in the standard text group, the regression coefficient *b* for number of misconceptions was -0.09 (SE = 0.05) and significantly different from zero, $t(52) = -1.85$, $p = .036$, $\Delta R^2 = .01$. Thus, an increase of one misconception led to a decrease of 9% in procedural text comprehension. In contrast, in the refutation text group, the regression coefficient for number of misconceptions was not significant, $b = 0.02$, $SE = 0.04$, $t(52) = 0.51$, $p = .611$, $\Delta R^2 = .01$, indicating that there was no significant effect of misconceptions.

Metacomprehension Accuracy

As shown in Table 4, the multiple regression with metacomprehension accuracy of conceptual comprehension revealed a marginal significant interaction effect between text type and misconceptions.

Table 4: Predictors of metacomprehension accuracy of conceptual comprehension.

Predictor	*b*	*SE b*	*t*(49)	*p*	ΔR^2
Constant	0.08	0.03	3.05	.004	
Text type	-0.01	0.05	-0.14	.892	<.01
NoM	0.02	0.02	0.84	.406	.01
Text type x NoM	-0.05	0.04	-1.35	.092	.04

Note. R^2 = .04, $F(3, 49) = 0.73$, $p = .541$. NoM = number of misconceptions.

Simple slopes analyses showed that, in the standard text group, the regression coefficient *b* for number of misconceptions was 0.04 (*SE* = 0.03) and marginally significantly different from zero, $t(52) = 1.43$, $p = .081$, ΔR^2 = .01. Hence, an increase of one misconception resulted in 4% greater overestimation of conceptual text comprehension. In contrast, in the refutation text group, the regression coefficient for number of misconceptions was not significant, $b = -0.01$, $SE = 0.02$, $t(52) = -0.38$, $p = .703$, $\Delta R^2 = .01$. Thus, there was no significant effect of misconceptions.

As can be seen in Table 5, the multiple regression with metacomprehension accuracy of procedural comprehension revealed no significant main effect or interaction effect.

Table 5: Predictors of metacomprehension accuracy of procedural comprehension.

Predictor	*b*	*SE b*	*t*(49)	*p*	ΔR^2
Constant	-0.11	0.05	-2.28	.027	
Text type	0.01	0.10	0.13	.897	<.01
NoM	0.03	0.03	1.05	.300	.02
Text type x NoM	-0.06	0.06	-0.94	.177	.02

Note. R^2 = .03, $F(3, 49) = 0.56$, $p = .643$. NoM = number of misconceptions.

Discussion

A large body of literature demonstrates positive learning effects from reading refutation texts compared with reading standard expository texts. However, little is known about whether refutation texts are also favorable for learning in statistics and for producing accurate self-assessments of comprehension. Thus, the present study is the first to address these questions.

First, the results showed that a refutational statistics text can compensate for the detrimental impact of misconceptions on text comprehension. When students read a standard text about covariance, a higher number of misconceptions about covariance led to poorer conceptual and procedural comprehension of the text. In contrast, when students read a refutation text about covariance, there was no effect of misconceptions on text comprehension. This result extends prior research by showing that refutation texts can prevent the adverse impact of misconceptions in statistics as well.

Importantly, the beneficial effect of the refutation text was demonstrated for both conceptual and procedural comprehension. Research in mathematics education widely acknowledges the view that conceptual and procedural knowledge are iteratively related to each other, with increases in conceptual knowledge leading to subsequent increases in procedural knowledge and vice versa (e.g., Rittle-Johnson & Schneider, 2015). Accordingly, in the present study, conceptual and procedural comprehension were quite strongly associated, $r = .46$, $p = .001$. Therefore, when students read the refutation text, their misconceptions were no longer predictive of both their acquisition of conceptual understanding and procedural skill.

Second, the findings showed that a refutational statistics text can compensate for the detrimental impact of misconceptions on metacomprehension accuracy with regard to conceptual comprehension. When students read a standard text about covariance, a higher number of misconceptions about covariance led to greater overestimation of conceptual comprehension. In contrast, when students read a refutation text about covariance, there was no significant effect of misconceptions on the accuracy with which they judged their conceptual comprehension. In accordance with the interpretation given by Prinz et al. (2017; see also Dunlosky et al., 2005), when reading a standard statistics text, students with a higher number of misconceptions likely constructed a flawed mental text representation. At the same time, when self-assessing their text comprehension, these students might have focused on the amount rather than on the correctness of the textual information they could access from memory. Accordingly, they might have more strongly overestimated their conceptual comprehension. However, when reading a refutational statistics text, the students might have been more inclined to assess the quality of the textual information they could retrieve from memory. This might have been the case because refutation texts promote the coactivation of misconceptions and the scientifically correct conceptions (van den Broek & Kendeou, 2008), thereby increasing the likelihood of knowledge evaluation and reflection in the context of conceptual change processes. Note, however, that the interaction effect between text type and misconceptions as well as the regression slope of misconceptions in the standard text group only approached the 10% level of statistical significance. As suggested by power analysis, this likely is the result of insufficient power to detect rather small effects. This is also supported by the findings of Prinz et al. (2017) that revealed a negative effect of misconceptions on metacomprehension accuracy of conceptual comprehension in the case of a standard text when using a sample of 49 participants. Therefore, future research should replicate the findings presented here while using larger sample sizes.

Contrary to expectation, however, the type of text and misconceptions did not affect metacomprehension accuracy of procedural comprehension. It can be assumed that the refutation text failed to coactivate procedural comprehension and, thus, decreased the likelihood that students would closely evaluate this type of comprehension. Yet, online measures such as think-aloud protocols could help to clarify the mechanisms proposed to underlie the effects observed in this study.

In sum, this study showed that refutation texts can compensate for detrimental effects of misconceptions on text comprehension and metacomprehension accuracy in the domain of statistics. Refutation texts appear to promote students to pay attention to inaccuracies in their knowledge, enhancing their self-regulated learning.

Acknowledgements

This research was supported by grants from the German Federal Ministry of Education and Research (01JA1518A).

References

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.

Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking.* Dordrecht, the Netherlands: Kluwer Academic.

Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* [Statistics for human and social scientists] (7th ed.). Berlin, Germany: Springer.

Cassidy, S. (2011). Self-regulated learning in higher education: Identifying key component processes. *Studies in Higher Education*, *36*, 989–1000.

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *International handbook of research on conceptual change*. New York, NY: Routledge.

Diakidoy, I.-A. N., Mouskounti, T., & Ioannides, C. (2011). Comprehension and learning from refutation and expository texts. *Reading Research Quarterly, 46*, 22–38.

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, *16*, 228–232.

Dunlosky, J., Rawson, K. A., & Middleton, E. L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language*, *52*, 551–565.

Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, *28*, 116–159.

Kendeou, P., & van den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology*, *97*, 235–245.

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

Liu, T.-C. (2010). Developing simulation-based computer assisted learning to correct students' statistical misconceptions based on cognitive conflict theory, using "correlation" as an example. *Journal of Educational Technology & Society*, *13*, 180–192.

Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology*, *97*, 723–731.

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

Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in a text: What readers believe versus what they read. *Psychological Science*, *3*, 229–235.

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

Prinz, A., Golke, S., & Wittwer, J. (2017). *The double curse of misconceptions: Misconceptions impair text comprehension and metacomprehension accuracy in the domain of statistics*. Manuscript submitted for publication.

Richter, T. (2007). How to analyze interactions of metric and categorical predictors: Not with median splits! *Zeitschrift für Medienpsychologie*, *19*, 116–125.

Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. In R. C. Kadosh & A. Dowker (Eds.), *The Oxford handbook of numerical cognition*. Oxford, UK: Oxford University Press.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*, 66–73.

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

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*, 335–351.

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.