Counteracting detrimental effects of misconceptions on learning and metacomprehension accuracy: The utility of refutation texts and think sheets

Anja Prinz¹ · Stefanie Golke¹ · Jörg Wittwer¹

Received: 25 February 2020 / Accepted: 11 February 2021 / Published online: 17 March 2021
© The Author(s) 2021

Abstract

Previous research has shown that misconceptions impair not only learners' text comprehension and knowledge transfer but also the accuracy with which they predict their comprehension and transfer. In the present experiment with $N=92$ university students, we investigated to what extent reading a refutation text or completing a think sheet compared with a control condition counteracts these adverse effects of misconceptions. The results revealed that both reading a refutation text and completing a think sheet supported learners with misconceptions in acquiring and accurately predicting their comprehension. Completing a think sheet additionally supported the learners in transferring their newly acquired knowledge, even though they were underconfident in their ability to do so. Moreover, learners who completed a think sheet were generally more accurate in discriminating between correctly and incorrectly answered test questions. Finally, delayed testing showed that the learning effects were quite stable, independent of the instructional method. This study reinforces the effectiveness of refutation texts and think sheets and provides important practical implications and avenues for future research.

Keywords Knowledge transfer · Metacomprehension accuracy · Misconceptions · Refutation texts · Text comprehension · Think sheets

A common challenge in instruction is that learners have misconceptions, which are incorrect beliefs that are typically well-embedded in their knowledge structures. When studying texts, learners with misconceptions not only achieve poorer text comprehension and knowledge transfer (e.g., Kendeou and Van den Broek 2005; Weingartner and Masnick 2019) but also more strongly overestimate their comprehension and transfer (Prinz et al. 2018, 2019). However, only when learners carefully monitor and accurately judge their learning from text, known as metacomprehension accuracy, can they engage in effective regulation processes to improve their performance (e.g., Rawson et al. 2011; Thiede et al. 2003; see, e.g., also Nelson and Narens 1990; Winne and Hadwin 1998). In the present study, we

✉ Anja Prinz
anja.prinz@ezw.uni-freiburg.de

¹ Department of Educational Science, University of Freiburg, Rempartstrasse 11, 79098 Freiburg, Germany

investigated instructional methods that might counteract the detrimental impact of misconceptions on learning (i.e., comprehension and transfer) and metacomprehension accuracy (i.e., prediction accuracy and response-confidence accuracy concerning comprehension and transfer). More precisely, we examined the utility of refutation texts, which are texts that explicitly state and refute misconceptions. In addition, we explored the effectiveness of think sheets, which are worksheets that require learners to extract and write down the correct information concerning misconceptions from a text.

Misconceptions impair learning from text and metacomprehension accuracy

A domain in which misconceptions are widespread is statistics (see, e.g., Huck 2016). Typical examples of statistical misconceptions refer to the concept of covariance. For instance, many learners erroneously believe that covariance proves a causal relationship between two variables or that covariance is a standardized statistic that takes on values within a finite range just like the correlation coefficient (e.g., Moritz 2004; Prinz et al. 2018, 2019). Such misconceptions have detrimental effects on learning from reading. Specifically, Prinz et al. (2019; see also Prinz et al. 2018) found that learners who had more misconceptions about covariance not only achieved poorer comprehension of a text on covariance but also provided more overconfident predictions of their comprehension than learners who had fewer misconceptions. The same result was found when the learners were required to use their acquired knowledge about covariance to understand an educational research report that presented empirical findings involving covariance: Compared with learners who had fewer misconceptions, learners who had more misconceptions revealed poorer knowledge transfer as well as more overconfident predictions of their transfer. These results can be explained by research and theories on text comprehension and metacomprehension. Deep and transferable comprehension requires learners to actively process the textual information and connect it with prior knowledge to construct a mental representation of the content—also known as the situation model (e.g., Kintsch 1998). The relevant prior knowledge becomes automatically activated in learners' memory. When learners' prior knowledge is contaminated with misconceptions, however, this can lead to an erroneous situation model (e.g., Kendeou and Van den Broek 2005; McNamara and Magliano 2009). Concerning metacomprehension, learners often use the amount of textual information they can access from memory as a cue for judging their comprehension (e.g., Dunlosky et al. 2005; see also the cue-utilization framework, Griffin et al. 2009; cf. Koriat 1997). When doing so, learners with misconceptions likely access a large proportion of incorrect information because their situation model is usually flawed. As a consequence, they tend to provide overconfident judgments.

The finding by Prinz et al. (2019) that misconceptions lead to overconfident judgments refers to *predictions*, which are prospective judgments made after reading but before testing. In addition to prospective judgements, retrospective judgments provided after testing can be relevant for effective self-regulated learning, such as when practice tests are applied. For example, *response-confidence judgments* follow the completion of each question in a test and indicate a learner's metacognitive awareness of the correctness of a given answer (e.g., Schraw 2009). Yet, the nature of monitoring captured by prospective and retrospective judgments differs. Prospective judgments are useful to assess monitoring during reading, whereas retrospective judgments are expedient to assess monitoring during testing (cf.,

e.g., Griffin et al. 2009, 2013). Research has shown that more accurate metacomprehension in terms of both predictions and response confidence leads to more adequate regulation (e.g., adaptive rereading decisions), which in turn results in enhanced performance (e.g., for predictions: De Bruin et al. 2011; Schleinschok et al. 2017; Thiede et al. 2003, 2012; for response confidence: Rawson et al. 2011). Therefore, work on instructional methods that support learners with misconceptions in learning and accurately judging their learning from reading is needed.

Refutation texts as a method to counteract detrimental effects of misconceptions

Refutation texts are texts that address misconceptions. To this end, a refutation-text passage typically contains three elements: (1) the statement of a misconception, (2) the alert that the misconception is false, and (3) the correct information (see, e.g., Braasch et al. 2013; Tippet 2010). Research has shown that refutation texts are typically more effective to improve learners' comprehension than standard texts that only provide the correct information (see, e.g., Guzzetti et al. 1993; Tippet 2010). Kendeou and colleagues proposed the *Knowledge Revision Components (KReC) framework* (Kendeou and O'Brien 2014) to explain how refutation texts promote knowledge revision, that is, conceptual change. The KReC framework outlines three necessary conditions for conceptual change. The central prerequisite for the revision of a misconception is its coactivation with the correct information in a learner's memory (*coactivation principle*). Because refutation texts report a misconception and the related correct information in close proximity, their coactivation is facilitated. Once coactivation has taken place, the new correct information needs to be integrated with prior knowledge, which ideally involves the transforming and outdateding of incorrect information (*integration principle*). As an increasing amount of correct information becomes integrated in the existent knowledge structures and is used, it draws more activation to itself and away from the misconception so that interference from the latter is reduced (*competing activation principle*).

Research on refutation texts has also shown that learners' improved comprehension after reading these texts is usually quite robust, even when tested after a time lag of several weeks or months (e.g., Cordova et al. 2014; Diakidoy et al. 2003; Donovan et al. 2018; Kendeou et al. 2014; Kowalski and Taylor 2017; Lassonde et al. 2016; Mason et al. 2008; Nussbaum et al. 2017). However, compared with immediate testing, the effect is often diminished at delayed testing, with learners partly reverting to the misconceptions (e.g., Cordova et al. 2014; Donovan et al. 2018; Kendeou et al. 2014; Kowalski and Taylor 2017; Lassonde et al. 2016; Mason et al. 2008; Nussbaum et al. 2017).

Furthermore, the study by Prinz et al. (2019) revealed that refutation texts can support metacomprehension accuracy. Specifically, reading a refutation text on covariance prevented the detrimental impact of misconceptions about covariance not only on learners' comprehension but also on their overconfidence when predicting their comprehension. Think-aloud data unveiled the mechanisms underlying these compensatory effects: When reading the refutation text, learners with more misconceptions more intensely engaged in conceptual-change processes, such as comparing information, experiencing cognitive conflict, and reacting to cognitive conflict, which in turn supported their comprehension and prediction accuracy. In line with prior research (e.g., Ariasi and Mason 2011, 2014; Diakidoy et al. 2011; Kendeou et al. 2011; Kendeou and Van den Broek 2007; McCrudden 2012;

McCrudden and Kendeou 2014; Van den Broek and Kendeou 2008) and the KReC framework (Kendeou and O'Brien 2014), it was assumed that reading the refutation text led to the coactivation of the misconceptions and the related correct information in the learners' memory, which then triggered further conceptual-change processes, such as contrasting and integrating information, and resulted in enhanced comprehension. Moreover, in the context of the conceptual-change processes, the learners seemed to have monitored and evaluated their understanding more carefully, thus providing more accurate predictions. Compensatory effects of reading the refutation text were also found when the learners had to transfer their knowledge about covariance to understand and judge their understanding of an educational research report. More precisely, reading the refutation text also prevented the detrimental impact of misconceptions on the learners' knowledge transfer and prediction accuracy of transfer (for transfer effects of refutation texts, see also Beker et al. 2019; Kendeou et al. 2016; Trevors et al. 2017; Weingartner and Masnick 2019).

Overall, the findings demonstrate that refutation texts benefit the comprehension and prediction accuracy of learners with misconceptions by inducing coactivation and conceptual-change processes and that these benefits transfer to situations where the acquired knowledge has to be put into practice (Prinz et al. 2019). Following these and prior results (e.g., Van den Broek and Kendeou 2008) as well as the KReC framework (Kendeou and O'Brien 2014), instructional methods that facilitate coactivation and conceptual-change processes should generally yield beneficial effects on learning and metacomprehension accuracy. To test this assumption, further instructional methods that have the potential to induce these processes should be explored. Besides theoretical insights, this would be important for several practical reasons. First, it can be challenging for instructors to appropriately construct refutation texts. Prior research has shown that several design features play a role for the effectiveness of refutation texts. For example, the effects of refutation texts on learning outcomes can vary depending on how explicitly a misconception is stated in text (Weingartner and Masnick 2019) and the source that the correct information is based on (e.g., a professor, blog, scientific article, encyclopedia; Kendeou et al. 2016; Trevors et al. 2017; Van Boekel et al. 2017). Second, an extensive use of refutation texts in instruction might be counterproductive because learners might experience a decline in motivation and become insensitive to the refutations. In fact, it is important to vary instructional methods to support learners' motivation and knowledge acquisition (see, e.g., Smith and Ragan 2005). Third, although refutation texts induce conceptual change, learners sometimes revert to misconceptions over time (e.g., Kendeou et al. 2014).

Think sheets as a method to counteract detrimental effects of misconceptions

Think sheets represent another instructional method to address misconceptions. A think sheet is a worksheet with questions about a topic and common incorrect beliefs about these questions. While reading a text on the topic, learners are required to extract and write down the correct information concerning the questions on the think sheet (Dole and Smith 1989; see also Brummer and Maccoca 2013; Fang et al. 2010; Fig. 2). In the study by Dole and Smith (1989), one group of fifth-grade students was taught a physics unit about matter by using think sheets in combination with the regular physics textbook. In the lessons, the students first wrote down their pre-existent beliefs about questions on a think sheet. Then, they recorded the correct answers to the questions based on the information in their textbook.

The other group of fifth-grade students was taught with regular textbook-based instruction only. The results showed that, after completion of the unit, the students in the think-sheet group held more scientifically correct conceptions about matter than the students in the control group. The superiority of the think-sheet group was still present when tested at an 8-week delay. Apart from the study by Dole and Smith (1989), however, there has been no systematic research on think sheets.

Theoretically, think sheets might facilitate similar cognitive processes as refutation texts, namely coactivation and conceptual-change processes. A reason for this assumption is that think sheets function as prompts as they instruct learners to contrast misconceptions with the correct information. As a consequence, the chances of coactivation and conceptual-change processes might be enhanced. In fact, prior research has shown that prompts, which are any questions or hints that indicate what knowledge or processes to use for productive learning, typically support learners' performance and transfer (e.g., Bannert et al. 2015; Berthold et al. 2007; Kramarski and Friedman 2014). In addition, completing a think sheet is a generative learning task. During generative learning, learners actively make sense of the material by (a) selecting the relevant information, (b) organizing it into a coherent structure, and (c) integrating it with their existing knowledge (see, e.g., Fiorella and Mayer 2016; Mayer 2009; Wittrock 1989). Accordingly, during think-sheet completion, learners need to filter the relevant correct information from a text, to mentally organize it, and to integrate it with their existent knowledge structures. When reading a refutation text, in contrast, the selection and organization processes are hardly required by the learners because the incorrect and correct information is presented directly one after another. Prior research has indicated that generative tasks enhance deep and transferable learning (see, e.g., Chi 2009; Fiorella and Mayer 2016). Hence, the generative task of completing a think sheet might provide additional benefits for comprehension and transfer. Moreover, prior research has shown that generative tasks, such as constructing concept maps and completing diagrams, increase learners' metacomprehension accuracy because they provide learners with valid cues about their deeper understanding (i.e., their situation model; see, e.g., Griffin et al. 2019; Prinz et al. 2020; Van de Pol et al. 2020). However, the generative tasks used in the prior research were not particularly designed to address misconceptions.

The present study

In the present study, we investigated to what extent reading a refutation text or completing a think sheet, in contrast to a control condition that involved reading a standard text together with a list of misconceptions, benefit the learning (i.e., comprehension and transfer) and metacomprehension accuracy (i.e., prediction accuracy and response-confidence accuracy concerning comprehension and transfer) of learners with misconceptions about the statistical concept of covariance. In addition, we explored how stable the potential benefits of the instructional methods are over time by applying a delayed posttest.

Prior studies have revealed that refutation texts (e.g., Kendeou and Van den Broek 2007) and think sheets (Dole and Smith 1989) support learning from reading. Both methods facilitate contrasting misconceptions with the correct information and might thus induce coactivation and conceptual-change processes. As a generative task, completing a think sheet might even do so in a more active way (see, e.g., Fiorella and Mayer 2016). Hence, we expected that, when reading the standard text, learners with a higher number of misconceptions would achieve poorer comprehension and transfer, whereas these negative effects

of misconceptions would not occur when reading the refutation text or completing the think sheet (*comprehension hypothesis* and *transfer hypothesis*). In addition, we explored whether these compensatory effects would be more pronounced with the think sheet.

Moreover, previous research has indicated that reading a refutation text can support metacomprehension accuracy by facilitating knowledge evaluation in the context of conceptual-change processes (Prinz et al. 2019). Completing a think sheet might evoke similar cognitive processes. Due to being a generative task, it might be even more beneficial by enhancing the access to cues related to the situation model (see, e.g., Griffin et al. 2019). Hence, we assumed that, when reading the standard text, learners with a higher number of misconceptions would show greater overconfidence in the predictions of their comprehension and transfer, as well as more inaccurate response confidence, but that these detrimental effects of misconceptions would not emerge when reading the refutation text or completing the think sheet (*comprehension-prediction hypothesis*, *transfer-prediction hypothesis*, and *response-confidence hypothesis*). Again, we also explored whether these compensatory effects would be more pronounced with the think sheet.

The positive effects of refutation texts (e.g., Kendeou et al. 2014) and think sheets (Dole and Smith 1989) on learning are usually quite persistent. However, while the effect of reading a refutation text sometimes diminishes over time (e.g., Cordova et al. 2014), the effect of the generative task of completing a think sheet might be more stable (see, e.g., Fiorella and Mayer 2016). Hence, we hypothesized that the expected results concerning comprehension and transfer would still exist when tested at a delay (*delayed-comprehension hypothesis* and *delayed-transfer hypothesis*).

Method

Participants

A total of $N=92$ students from a German university participated in this study. They were on average 23.32 ($SD=2.61$) years old, with 66% being female and 34% being male. Students were required to have attended at least one statistics course that covered the topic of covariance. Thus, they had some prior knowledge about the topic and possibly misconceptions. On average, the participants were in their 4.13 ($SD=2.06$) semester and had attended 1.57 ($SD=0.80$) statistics courses that covered covariance. They studied educational science, cognitive science, psychology, economics, or environmental science. They received course credit or a financial reward for taking part in the study.

Design

The study had an experimental, between-subjects, design with two independent variables, instructional method and misconceptions. Concerning instructional method, the participants were randomly assigned to the refutation-text group ($n=31$), the think-sheet group ($n=29$), or the control group ($n=32$). Misconceptions were not experimentally manipulated but measured.

The dependent variables were comprehension, knowledge transfer, prediction accuracy of comprehension, prediction accuracy of transfer, and response-confidence accuracy. Comprehension and transfer were assessed twice: directly after studying and a week later.

<p>(a)</p> <p>... As a nonstandardized statistic, covariance is dependent on the measuring units of the underlying variables. Theoretically, covariance values can reach from negative infinity to positive infinity. For example, if in two studies the variables weight and/or height are measured in different units (e.g., weight: kg, g; height: m, cm), the studies will yield covariances between these two variables that are not directly comparable. ...</p>
<p>(b)</p> <p>... (1) Learners often assume that, like the correlation coefficient, covariance is a standardized statistic that always takes on values within a finite range (-1 to +1), enabling direct comparisons of different covariance values. (2) However, this assumption is incorrect. (3) As a nonstandardized statistic, covariance is dependent on the measuring units of the underlying variables. Theoretically, covariance values can reach from negative infinity to positive infinity. For example, if in two studies the variables weight and/or height are measured in different units (e.g., weight: kg, g; height: m, cm), the studies will yield covariances between these two variables that are not directly comparable. ...</p>

Fig. 1 Sample excerpts from **(a)** the standard text and **(b)** the refutation text with indication of the elements of a refutation-text passage (not provided in the original text; translated from German). 1 = description of the misconception; 2 = statement of its incorrectness; 3 = correct information

Materials

Statistics text

The statistics text on covariance was adapted from a commonly used German statistics textbook by Bortz and Schuster (2010; cf. Prinz et al. 2018, 2019). The text described conceptual aspects of covariance, such as its different directions, and explained procedural aspects of covariance, such as its calculation. The text was modified so that two versions existed: a standard- and a refutation-text version.

The refutation text challenged four common misconceptions about covariance (e.g., Batanero et al. 1996; Liu et al. 2009; Prinz et al. 2018, 2019): (1) covariance implies causality, (2) covariance is a standardized statistic, (3) covariance is related to the slope of the fit line, and (4) zero covariance proves the absence of any association. For each misconception, the three typical elements of a refutation-text passage were provided (see, e.g., Tippett 2010): First, one sentence described the misconception. Second, one sentence explicitly stated the incorrectness of the misconception. Third, three sentences provided the correct information. In contrast to the refutation text, the standard text only provided the correct information concerning each misconception (see Fig. 1 for an example). The standard text included 536 words (27 sentences) and had a Flesch-Reading-Ease score of 33, whereas the refutation text included 665 words (35 sentences) and had a Flesch-Reading-Ease score of 34 (Flesch 1948). The two text versions were not adjusted in length to avoid confounding the manipulation with other variations (e.g., inclusion of additional or repetitive information in the standard text or omission of information from the refutation text). By definition, a refutation text is longer than a standard text because it states misconceptions in addition to the correct information (cf., e.g., Diakidoy et al. 2011; Mikkilä-Erdmann 2001).

Central question about covariance	Typical misconception concerning the question	Correct explanation concerning the question
What does covariance reveal about a cause-and-effect relationship between two variables?	Covariance between two variables indicates that one variable causally affects the other variable so that one can infer a cause-and-effect relationship.	
How can the magnitude of covariance values be interpreted?	Like the correlation coefficient, covariance is a standardized statistic that always takes on values within a finite range (-1 to +1), enabling direct comparisons of different covariance values.	
How can the strength of the covariance between two variables be inferred from a scatterplot?	The steeper the line in the scatterplot that represents the covariance between two variables, the stronger the covariance.	
What does a covariance of zero between two variables mean?	A covariance of zero proves the absence of any kind of relationship between two variables.	

Fig. 2 The think sheet used in the present study (translated from German)

Think sheet

The think sheet contained a table with three columns (see Fig. 2). The first column presented four questions concerning covariance. In the second column, a typical misconception concerning each question was described. These descriptions of the misconceptions were equivalent to the sentences in the refutation text that outlined the four common misconceptions about covariance (first sentence of each refutation-text passage). The third column was left blank. The participants in the think-sheet group were required to fill out this column with the correct information concerning the questions provided in the text on covariance (standard-text version).

Control sheet

A control sheet that contained only the first two columns (i.e., a column with questions and a column with the respective misconceptions) of the table on the think sheet was used in

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

- ☐ The two sports scientists received the same covariance because it does not matter that sports scientist A used seconds and sports scientist B used milliseconds for the calculation. (*misconception*)
- ☐ The calculations of both sports scientists yielded no covariance because one cannot calculate covariance from time data. (*incorrect*)
- ☐ No statement about the two covariances can be made because it is unknown whether the time and height variables were linear. (*incorrect*)
- ☐ Sports scientist B obtained a higher covariance than sports scientist A because milliseconds yield bigger numbers than seconds. (*correct*)

Fig. 3 Sample question of the misconceptions test (translated from German)

the control group. This was done to keep the amount of information (i.e., the learning content) presented to the participants in all groups constant.

Measures

Misconceptions

The misconceptions test consisted of four questions that assessed the four common misconceptions about covariance (see the [Materials](#) section on the statistics text; cf. Prinz et al. 2018, 2019). The questions had a single-choice format with four response options. One option represented the particular misconception, one option represented the correct answer, and the two remaining options represented incorrect answers (see Fig. 3 for an example). Thus, a participant could record between 0 and 4 misconceptions. Because the misconceptions referred to multiple, rather distinct aspects of covariance, we refrained from calculating internal consistency.

Comprehension

The four questions of the misconceptions test were adopted to assess comprehension of the statistics text. The participants received 1 point for each correct answer. Thus, they could achieve a maximum of 4 points.

Transfer

Transfer was measured by assessing comprehension of an educational research report on scientific literacy adapted from PISA (Organisation for Economic Co-operation and Development 2006). This text described findings about the covariance between aspects of scientific literacy (e.g., familiarity with environmental issues) and academic

... Besides the INSA study, there are more studies that dealt with scientific literacy. For example, in a study in Sweden, it was investigated whether there is a relationship between paternal gross income in Swedish krona and students' familiarity with environmental issues. The study revealed a relatively high covariance of 580. A similar study was conducted with students in Germany. This study also revealed a covariance of 580 between paternal gross income in Euro and students' familiarity with environmental issues. ...

Fig. 4 Sample excerpt from the educational research report (translated from German)

A study in Sweden revealed a covariance of 580 between paternal gross income in Swedish krona and students' familiarity with environmental issues. Likewise, a similar study in Germany found a covariance of 580 between paternal gross income in Euro and students' familiarity with environmental issues. Which statement concerning the results of the two studies can be made?

- ☐ Students in the two countries get familiarized with environmental issues to the same degree. (*incorrect*)
- ☐ The strength of the relationship between paternal gross income and students' familiarity with environmental issues is equal in the two countries. (*misconception*)
- ☐ Students in the two countries get familiarized with environmental issues to a different degree. (*incorrect*)
- ☐ The strength of the relationship between paternal gross income and students' familiarity with environmental issues is not equal in the two countries. (*correct*)

Fig. 5 Sample question of the transfer test (translated from German)

achievement in science. Comprehension of this text represented transfer because understanding it required the participants to transfer their previously acquired knowledge about covariance to a new context, namely studying another text, and to a different content domain, namely the domain of educational science. In addition, the participants had to spontaneously apply the learned information about covariance to the distant domain as there were no hints or prompts to do so (see Barnett and Ceci 2002). Specifically, four passages of the educational research report presented information that learners with particular misconceptions about covariance were likely to misunderstand. These four passages tapped the same four misconceptions as the statistics text (see Fig. 4 for an example). The text comprised 655 words and had a Flesch-Reading-Ease score of 25 (Flesch 1948).

The four transfer questions addressed the same four misconceptions about covariance as the comprehension questions. However, the questions were altered and embedded in different cover stories to avoid memory and fatigue effects due to multiple exposure to the same questions (see Fig. 5 for an example). The participants received 1 point for each correct answer. Thus, they could achieve a maximum of 4 points.

Prediction accuracy

Prediction accuracy was operationalized as *bias*. After having read the statistics text, the participants predicted the number of questions they would presumably be able to answer correctly. They made their predictions separately for the comprehension and transfer questions (“Of the 4 comprehension/transfer questions, I will answer __ questions correctly.”). To do so, they were informed that the comprehension questions would tap their understanding of the presented concepts, whereas the transfer questions would require them to apply their acquired understanding. Bias was computed as the signed difference between a participant’s predicted and actual number of questions correct (see, e.g., Schraw 2009). Thus, a positive value indicated overconfidence, a negative value underconfidence, and a value of zero a perfectly accurate judgment.

Response-confidence accuracy

Response-confidence accuracy was operationalized as *relative accuracy*. After the completion of each test question, the participants indicated how confident they were that their answer was correct on a Likert scale from 0 (*very uncertain*) to 5 (*very certain*). Relative accuracy was computed as the gamma correlation between a participant’s response-confidence judgments and actual performance scores across the set of questions (see, e.g., Schraw 2009). Calculating this measure requires some variation within a participant’s judgments and performance. Thus, we computed the correlation across all eight test questions instead of across the comprehension and transfer questions separately. A higher positive value indicated higher accuracy because a participant was generally confident in correct answers and unconfident in incorrect answers.

Procedure

The experiment comprised two sessions: an immediate and a delayed session. During the immediate session, the participants first completed the misconceptions test. After a short filler task, they read the statistics text on covariance. The participants in the *refutation-text group* received the refutation-text version of the statistics text. The time for reading was set to 12 min. The participants in the *think-sheet group* were first provided with the think sheet. After studying it for 2 min, they were additionally provided with the standard-text version of the statistics text. They were instructed to complete the third column of the table on the think sheet by answering the questions with the correct information provided in the text. They were told that they should answer each question as concretely and completely as possible. For this task, additional 10 min were assigned. The participants in the *control group* first received the control sheet and were given 2 min for studying it. Thereafter, they were additionally provided with the standard-text version of the statistics text. They were not required to complete any task besides reading for which they had further 10 min. Hence, the overall study time was kept constant across the groups. Moreover, all participants were instructed to read carefully and were informed that their understanding would be tested. After studying, the participants made the predictions of their comprehension and transfer performance. To do so, they were informed about the nature of the questions, the question format, and the number of questions. Then, the participants answered the comprehension questions and indicated their confidence in each response. Thereafter, they read the

educational research report within 8 min, answered the transfer questions, and rated their confidence in each response. Finally, the participants answered demographic questions. All materials except the think and control sheets were presented on a computer. This session took approximately 75 min. In the delayed session, which took place 6 to 8 days later and lasted about 15 min, the participants again answered the comprehension and transfer questions.

Analysis

To examine our hypotheses, we conducted multiple regression analyses with misconceptions as the predictor variable and instructional method as the moderator variable following the approach by Hayes (2018) and using the PROCESS macro (Version 3.4) for SPSS. Because instructional method was a multicategorical moderator with three groups, two variables coding group membership were generated. To do so, indicator coding was used, with the control group being the reference group. Hence, the variable refutation-text/control was set to 1 for participants in the refutation-text group and 0 otherwise, and the variable think-sheet/control was set to 1 for participants in the think-sheet group and 0 otherwise (cf. Hayes 2018). The predictor misconceptions was mean centered to obtain meaningful estimates of the effects of the instructional methods. More precisely, by applying mean centering, the regression coefficients for the instructional methods estimate the difference in the dependent variable between the control group and the refutation-text group (refutation-text/control) or the think-sheet group (think-sheet/control) at an average number of misconceptions. In each regression model, misconceptions, the two variables coding group membership, and two interaction terms between misconceptions and the grouping variables were included (cf. Hayes 2018).

A hypothesis about moderation of the effect of a predictor on a dependent variable by a multicategorical moderator can be tested by comparing the fit of two models: The first model constrains the predictor's effect on the dependent variable to be linearly independent of the moderator. The second model allows the predictor's effect to linearly depend on the moderator. If the second model fits significantly better than the first model, this provides evidence that the predictor interacts with the moderator. Hence, in the present analyses, if there is an interaction between misconceptions and instructional method, the model should fit better when the interaction terms are included compared with when they are not. This approach involves constructing the difference in R^2 between the two models (i.e., ΔR^2) and converting this difference into an F -ratio. To further probe a significant overall interaction, we quantified the relationship between the predictor and the dependent variable in each of the three groups and looked at the difference between these conditional effects (cf. Hayes 2018).

Results

The descriptive statistics and intercorrelations among the variables of this study are presented in Table 1. Moreover, Table 2 shows the descriptive statistics for the three groups separately. One participant did not complete the delayed session of the experiment. Consequently, only 91 participants were included in the respective analyses. An alpha level of .05 was used for all analyses.

Table 1 Descriptive statistics and intercorrelations among the variables

Measure	1	2	3	4	5	6	7	8	9	10
1. Misconceptions	–									
2. Comprehension	.01	–								
3. Transfer	.23*	.52*	–							
4. Bias of comprehension predictions	.11	–.77*	–.28*	–						
5. Bias of transfer predictions	–.03	–.29*	–.76*	.31*	–					
6. Response-confidence accuracy	–.14	.33*	.05	–.23*	.03	–				
7. Delayed comprehension ^a	.05	.79*	.49*	–.47*	–.18	.27*	–			
8. Delayed transfer ^a	.26*	.60*	.81*	–.29*	–.48*	.06	.65*	–		
9. Number of statistics courses on covariance	.10	–.14	.01	.10	.01	–.11	–.07	.02	–	
10. Number of semester	–.04	.10	.08	–.02	–.07	.01	.15	.10	.16	–
<i>M</i>	1.88	2.85	2.57	0.26	–0.22	.50	2.73	2.66	1.57	4.13
<i>SD</i>	0.96	1.10	1.12	1.18	1.19	.60	1.06	1.14	0.80	2.06

* $p < .05$ ^aOne participant did not complete the delayed session. Thus, only $N = 91$ participants were included in the respective analyses

Table 2 Means and standard deviations of the variables in the three groups

Variable	Refutation-text group (<i>n</i> = 31)		Think-sheet group ^a (<i>n</i> = 29)		Control group (<i>n</i> = 32)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Misconceptions	2.03	0.95	1.79	0.98	1.81	0.96
Comprehension	3.10	0.79	2.69	1.14	2.75	1.30
Transfer	2.71	0.98	2.59	1.21	2.41	1.19
Bias of comprehension predictions	0.07	0.93	0.41	0.98	0.31	1.51
Bias of transfer predictions	−0.32	1.08	−0.17	1.14	−0.16	1.37
Response-confidence accuracy ^b	.43	.60	.72	.35	.38	.74
Delayed comprehension	2.94	0.85	2.57	1.03	2.66	1.23
Delayed transfer	2.84	1.13	2.54	1.23	2.59	1.07
Number of statistics courses on covariance	1.53	0.79	1.59	0.68	1.58	0.93
Number of semester	4.06	1.93	4.38	2.23	3.97	2.07

^aIn this group, one participant did not complete the delayed session. Thus, only *n* = 28 participants were included in the respective analyses

^bDue to invariance in performance scores, Gamma could not be computed for three participants in the refutation-text group, six participants in the think-sheet group, and five participants in the control group

Preliminary analyses

A priori group comparisons

The three groups did not significantly differ from each other with regard to the number of semesters, $F(2, 89) = 0.32$, $p = .726$, $\eta^2 = .01$ (small effect), the number of attended statistics courses on covariance, $F(2, 89) = 0.04$, $p = .961$, $\eta^2 < .01$ (small effect), and the number of misconceptions prior to reading, $F(2, 89) = 0.58$, $p = .560$, $\eta^2 = .01$ (small effect; see Table 2 for the descriptive statistics). In addition, the number of semesters and the number of statistics courses on covariance did not significantly correlate with the independent and dependent variables. Therefore, they were not considered as covariates.

Reduction of misconceptions in the groups

To examine to what extent the instructional methods reduced misconceptions, we performed a mixed ANOVA with instructional method (refutation text, think sheet, and control) as the between-subjects variable and time of misconceptions assessment (misconceptions test before studying vs. comprehension test in the immediate session vs. comprehension test in the delayed session) as the within-subject variable. To do so, we determined how many misconceptions the participants held before studying, at immediate testing, and at delayed testing (see Fig. 6). The assumption of homogeneity of variances was fulfilled. However, the Mauchly's test indicated violation of the assumption of sphericity, $\chi^2(2) = 27.75$, $p < .001$. Thus, we adjusted the results by using the Huynh–Feldt correction. The ANOVA showed that there was no significant main effect of instructional method,

Fig. 6 Reduction of misconceptions in the three groups. In the think-sheet group, one participant did not complete the delayed session. Thus, this repeated measures analysis was based on only $n=28$ participants of this group

$F(2, 88)=0.69$, $p=.507$, $\eta^2=.02$ (small effect). Yet, there was a significant main effect of time of misconceptions assessment, $F(1.63, 143.52)=44.59$, $p<.001$, $\eta^2=.34$ (large effect). This main effect was qualified by a marginally significant interaction effect between instructional method and time of misconceptions assessment, $F(3.26, 143.52)=2.22$, $p=.083$, $\eta^2=.05$ (medium effect).

A simple effects analysis with Bonferroni correction of the alpha level demonstrated that, compared to before studying, the number of misconceptions was reduced at immediate testing and delayed testing with all three instructional methods ($ps \leq .010$). Moreover, this reduction in misconceptions was relatively stable. That is, for the think-sheet group and the control group, there was no significant change in the number of misconceptions from immediate to delayed testing ($ps=1.00$). For the refutation-text group, however, there was a significant increase in misconceptions from immediate to delayed testing ($p=.045$); although the number of misconceptions was still lower than in the other two groups. Hence, it can be concluded that all three instructional methods were effective in inducing conceptual change.¹

Quality of completed think sheets

As a treatment check, we examined the quality of the participants' answers provided on the think sheet. An answer was scored with 1 when it was correct and complete, and with 0 when it was incorrect or incomplete. The results showed that 23 of the 29 participants achieved the maximum score of 4 points, four participants achieved 3 points, and two participants achieved 2 points. Hence, the participants completed the think sheets thoroughly.

¹ Note also that a further simple effects analysis with Bonferroni correction of the alpha level showed that, at immediate testing, the refutation-text group had marginally significantly fewer misconceptions than the control group ($p=.077$) and the think-sheet group ($p=.093$). At delayed testing, there was no significant difference between the groups with regard to the number of misconceptions anymore ($ps \geq .956$).

Fig. 7 Effect of misconceptions on comprehension as a function of instructional method

Table 3 Multiple regression analysis on comprehension

Predictor variable	<i>b</i>	<i>SE</i>	<i>t</i> (86)	<i>p</i>
Constant	2.72	0.19	14.26	<.001
Misconceptions	−0.40	0.20	−1.99	.050
Refutation-text/control	0.35	0.28	1.29	.202
Think-sheet/control	−0.01	0.28	−0.04	.971
Misconceptions × refutation-text/control	0.54	0.29	1.88	.063
Misconceptions × think-sheet/control	0.66	0.29	2.30	.024

$$R^2 = .09, F(5, 86) = 1.72, p = .138$$

Comprehension

The analysis on comprehension revealed that there was a marginally significant overall interaction between misconceptions and instructional method, $\Delta R^2 = .06$, $F(2, 86) = 3.03$, $p = .053$. In the control group, misconceptions had a marginally significant negative effect on comprehension, $b = -0.40$, $SE = 0.20$, $t(91) = -1.99$, $p = .050$. In contrast, misconceptions had no significant effect in the refutation-text group, $b = 0.15$, $SE = 0.21$, $t(91) = 0.70$, $p = .487$, and the think-sheet group, $b = 0.27$, $SE = 0.21$, $t(91) = 1.28$, $p = .204$. Figure 7 provides a graphical representation of the conditional effects.

As can be seen in Table 3, the interaction between misconceptions and refutation-text/control was marginally significant. Hence, the difference between the impact of misconceptions on comprehension in the control group and the refutation-text group (which corresponds to the difference between the slopes of the lines for the two groups) was marginally significant. Moreover, the interaction between misconceptions and think-sheet/control was significant. Thus, the impact of misconceptions on comprehension in the control group and the think-sheet group differed significantly.

Fig. 8 Effect of misconceptions on transfer as a function of instructional method

We further examined whether the impact of misconceptions on comprehension differed between the refutation-text group and the think-sheet group. To this end, the refutation-text group was coded as the reference group and the analysis was run again. The results showed that the interaction between misconceptions and think-sheet/refutation-text was not significant, $b=0.12$, $SE=0.21$, $t(86)=0.42$, $p=.679$.

In sum, the results were in line with the comprehension hypothesis in that the refutation text and the think sheet compensated the negative impact of misconceptions on comprehension occurring in the control condition. This compensatory effect did not differ between the refutation text and the think sheet.

Transfer

The analysis on transfer showed that there was a significant overall interaction between misconceptions and instructional method, $\Delta R^2=.08$, $F(2, 86)=4.05$, $p=.021$. The effect of misconceptions on transfer was not significant in the control group, $b=-0.16$, $SE=0.20$, $t(91)=-0.79$, $p=.430$, and the refutation-text group, $b=0.31$, $SE=0.21$, $t(91)=1.49$, $p=.139$. However, in the think-sheet group, misconceptions had a significant positive effect, $b=0.66$, $SE=0.21$, $t(91)=3.17$, $p=.002$ (see Fig. 8).

The interaction between misconceptions and refutation-text/control was not significant. Yet, the interaction between misconceptions and think-sheet/control was significant. Hence, the impact of misconceptions on transfer in the control group and the think-sheet group differed significantly (see Table 4).

We further investigated whether the impact of misconceptions on transfer differed between the refutation-text group (reference group) and the think-sheet group. However, the interaction between misconceptions and think-sheet/refutation-text was not significant, $b=0.35$, $SE=0.29$, $t(86)=1.19$, $p=.238$.

Taken together, the results were not in line with the transfer hypothesis. Misconceptions had no significant negative effect on transfer in the control group. Consequently, no compensatory effect of the refutation text and the think sheet could emerge. However,

Table 4 Multiple regression analysis on transfer

Predictor variable	<i>b</i>	<i>SE</i>	<i>t</i> (86)	<i>p</i>
Constant	2.40	0.19	12.64	< .001
Misconceptions	−0.16	0.20	−0.79	.430
Refutation-text/control	0.27	0.27	0.99	.328
Think-sheet/control	0.25	0.28	0.90	.370
Misconceptions × refutation-text/control	0.47	0.29	1.63	.108
Misconceptions × think-sheet/control	0.81	0.29	2.83	.006

$$R^2 = .14, F(5, 86) = 2.83, p = .020$$

Fig. 9 Effect of misconceptions on bias of comprehension predictions as a function of instructional method

unexpectedly, in the think-sheet group, participants who had more misconceptions in the beginning were even better able to transfer their newly acquired comprehension.

Bias of comprehension predictions

The analysis on bias of comprehension predictions revealed that there was a marginally significant overall interaction between misconceptions and instructional method, $\Delta R^2 = .06$, $F(2, 86) = 2.76$, $p = .069$. In the control group, the effect of misconceptions on bias was significant and positive, $b = 0.55$, $SE = 0.22$, $t(91) = 2.56$, $p = .012$, indicating that more misconceptions resulted in greater overconfidence. In contrast, misconceptions had no significant effect in the refutation-text group, $b < -0.01$, $SE = 0.22$, $t(91) = -0.01$, $p = .991$, and the think-sheet group, $b = -0.13$, $SE = 0.22$, $t(91) = -0.59$, $p = .557$ (see Fig. 9).

The interaction between misconceptions and refutation-text/control was marginally significant. Therefore, the impact of misconceptions on bias of comprehension predictions in the control group and the refutation-text group differed marginally significantly. In addition, the interaction between misconceptions and think-sheet/control was significant. Thus, the impact of misconceptions on bias of comprehension predictions in the control group and the think-sheet group differed significantly (see Table 5).

Table 5 Multiple regression analysis on bias of comprehension predictions

Predictor variable	<i>b</i>	<i>SE</i>	<i>t</i> (86)	<i>p</i>
Constant	0.35	0.21	1.71	.091
Misconceptions	0.55	0.22	2.56	.012
Refutation-text/control	−0.29	0.29	−0.97	.334
Think-sheet/control	0.05	0.30	0.18	.860
Misconceptions × refutation-text/control	−0.55	0.31	−1.79	.078
Misconceptions × think-sheet/control	−0.68	0.31	−2.20	.031

$$R^2 = .09, F(5, 86) = 1.67, p = .150$$

Fig. 10 Effect of misconceptions on bias of transfer predictions as a function of instructional method

We further analyzed whether the impact of misconceptions on bias of comprehension predictions differed between the refutation-text group (reference group) and the think-sheet group but found no significant interaction, $b = -0.13$, $SE = 0.22$, $t(86) = -0.59$, $p = .557$.

Overall, the results were in line with the comprehension-prediction hypothesis in that the refutation text and the think sheet counteracted the detrimental impact of misconceptions on the accuracy of comprehension predictions occurring in the control condition. This compensatory effect did not differ between the refutation text and the think sheet.

Bias of transfer predictions

The analysis on bias of transfer predictions revealed that there was a significant overall interaction between misconceptions and instructional method, $\Delta R^2 = .07$, $F(2, 86) = 3.17$, $p = .047$. The effect of misconceptions on bias was not significant in the control group, $b = 0.35$, $SE = 0.22$, $t(91) = 1.58$, $p = .117$, and the refutation-text group, $b = -0.03$, $SE = 0.23$, $t(91) = -0.11$, $p = .912$. Yet, there was a marginally significant negative effect of misconceptions in the think-sheet group, $b = -0.45$, $SE = 0.23$, $t(91) = -1.97$, $p = .052$. More misconceptions resulted in more underconfident predictions (see Fig. 10).

The interaction between misconceptions and refutation-text/control was not significant. In contrast, the interaction between misconceptions and think-sheet/control was significant.

Table 6 Multiple regression analysis on bias of transfer predictions

Predictor variable	<i>b</i>	<i>SE</i>	<i>t</i> (86)	<i>p</i>
Constant	−0.13	0.21	−0.63	.529
Misconceptions	0.35	0.23	1.58	.117
Refutation-text/control	−0.19	0.30	−0.62	.537
Think-sheet/control	−0.08	0.30	−0.26	.795
Misconceptions × refutation-text/control	−0.37	0.32	−1.18	.241
Misconceptions × think-sheet/control	−0.80	0.32	−2.52	.014

$$R^2 = .07, F(5, 86) = 1.35, p = .250$$

Table 7 Multiple regression analysis on response-confidence accuracy

Predictor variable	<i>b</i>	<i>SE</i>	<i>t</i> (72)	<i>p</i>
Constant	0.38	0.12	3.35	.001
Misconceptions	−0.18	0.13	−1.35	.182
Refutation-text/control	0.05	0.16	0.30	.768
Think-sheet/control	0.33	0.17	1.89	.062
Misconceptions × refutation-text/control	0.15	0.18	0.87	.387
Misconceptions × think-sheet/control	0.14	0.19	0.76	.453

$$R^2 = .08, F(5, 72) = 1.28, p = .284$$

Thus, the impact of misconceptions on bias of transfer predictions in the control group and the think-sheet group differed significantly (see Table 6).

We further examined whether the impact of misconceptions on bias of transfer predictions differed between the refutation-text group (reference group) and the think-sheet group but found no significant interaction, $b = -0.43$, $SE = 0.32$, $t(86) = -1.32$, $p = .192$.

Taken together, the results were not in line with the transfer-prediction hypothesis. Misconceptions had no significant influence on bias of transfer predictions in the control group so that no compensatory effect of the refutation text and the think sheet could emerge. In addition, unexpectedly, in the think-sheet group, participants who started with more misconceptions were more underconfident in their ability to transfer their acquired comprehension.

Response-confidence accuracy

We could not calculate response-confidence accuracy for 14 participants because there was no variance in their performance scores. The analysis revealed that there was no significant overall interaction between misconceptions and instructional method, $\Delta R^2 = .01$, $F(2, 72) = 0.44$, $p = .646$. As shown in Table 7, none of the two interaction terms was significant. However, the effect of think-sheet/control was marginally significant, indicating that response-confidence accuracy was higher in the think-sheet group than in the control group at an average number of misconceptions (because mean centering was applied).

Table 8 Multiple regression analysis on delayed comprehension

Predictor variable	<i>b</i>	<i>SE</i>	<i>t</i> (85)	<i>p</i>
Constant	2.65	0.19	14.07	< .001
Misconceptions	−0.18	0.20	−0.89	.377
Refutation-text/control	0.26	0.27	0.97	.333
Think-sheet/control	−0.06	0.28	−0.22	.825
Misconceptions × refutation-text/control	0.36	0.28	1.28	.205
Misconceptions × think-sheet/control	0.30	0.29	1.07	.289

$R^2 = .05$, $F(5, 85) = 0.79$, $p = .558$

Given this finding, we further explored the overall difference between the groups with regard to response-confidence accuracy. Specifically, we examined whether response-confidence judgments were more accurate in the think-sheet group than in the other two groups by conducting a contrast analysis with the following contrast weights: think-sheet group: +2, refutation-text group: −1, control group: −1. The contrast turned out to be significant, $F(1, 75) = 4.43$, $p = .039$, $\eta^2 = .06$ (medium effect).

As misconceptions and instructional method did not interact to influence response-confidence accuracy, the response-confidence hypothesis was not supported. It appeared, however, that response-confidence accuracy was higher in the think-sheet group than in the other two groups.

Delayed comprehension

The analysis on delayed comprehension showed that there was no significant overall interaction between misconceptions and instructional method, $\Delta R^2 = .01$, $F(2, 85) = 0.95$, $p = .391$. In addition, none of the variables entered in the regression analysis were significant (see Table 8). Hence, the results appeared to be not in line with the delayed-comprehension hypothesis.

Delayed transfer

The analysis on delayed transfer revealed that there was no significant overall interaction between misconceptions and instructional method, $\Delta R^2 = .04$, $F(2, 85) = 1.69$, $p = .190$. Moreover, none of the variables entered in the regression analysis were significant (see Table 9). Thus, the delayed-transfer hypothesis was apparently not confirmed.

Table 9 Multiple regression analysis on delayed transfer

Predictor variable	<i>b</i>	<i>SE</i>	<i>t</i> (85)	<i>p</i>
Constant	2.59	0.20	13.24	< .001
Misconceptions	−0.02	0.21	−0.07	.941
Refutation-text/control	0.18	0.28	0.63	.530
Think-sheet/control	−0.01	0.29	−0.05	.960
Misconceptions × refutation-text/control	0.47	0.30	1.58	.119
Misconceptions × think-sheet/control	0.47	0.30	1.59	.115

$$R^2 = .11, F(5, 85) = 2.07, p = .077$$

Discussion

In this study, we investigated the effectiveness of reading a refutation text or completing a think sheet in comparison to reading a standard text together with a list of misconceptions. Specifically, we examined the extent to which these instructional methods support learners who have misconceptions about the statistical concept of covariance in learning (i.e. comprehension and transfer) and accurately judging their learning (i.e., prediction accuracy and response-confidence accuracy concerning comprehension and transfer).

Effects of misconceptions and instructional method on comprehension

The results revealed that learners with more misconceptions achieved poorer comprehension than learners with fewer misconceptions when reading the standard text. However, this negative effect of misconceptions was compensated for when learners read the refutation text or completed the think sheet. This finding is in line with prior research showing that refutation texts and think sheets can prevent the adverse impact of misconceptions on comprehension (e.g., Dole and Smith 1989; Kendeou and Van den Broek 2007). Although, in contrast to reading a refutation text, completing a think sheet is a generative task (see, e.g., Fiorella and Mayer 2016), the compensatory effects of the two instructional methods did not differ from each other. The two methods might induce similar cognitive processes. By reporting a misconception and the related correct information one after another, refutation texts induce coactivation and conceptual-change processes that result in an improved comprehension (e.g., Kendeou and Van den Broek 2007; Prinz et al. 2019). By prompting learners to extract the correct information from a text and to set it against the respective misconception, think sheets likely evoke coactivation and conceptual-change processes as well. Thus, the present findings are in line with the KReC framework that emphasizes the importance of coactivation and conceptual-change processes for enhancing the comprehension of learners whose knowledge is contaminated with misconceptions (Kendeou and O'Brien 2014; see, e.g., also Van den Broek and Kendeou 2008). Prior studies on the KReC framework focused on refutation texts. Hence, by showing that also other instructional methods that induce the relevant processes are effective, the present study extends this research and increases confidence in the assumptions made by the KReC

framework, especially the coactivation principle (Kendeou and O'Brien 2014). It has no additional benefit, however, when the processes are triggered within a generative task such as completing a think sheet that requires the selection and organization of the information besides the contrast and integration. Apparently, the key to foster coactivation and conceptual-change processes is that the direct contrast between misconceptions and the correct information is facilitated and it plays no role how this is achieved. Alternatively, the deep learning processes and the search processes occurring during think-sheet completion might cancel each other out. More specifically, although the generative task promotes elaborative learning processes, it also requires search processes that might induce a high cognitive load (i.e., extraneous load; e.g., Chandler and Sweller 1991). That is, learners have to devote working-memory capacity to filtering and organizing the relevant information, which thus cannot be used for comprehending the information or monitoring.

Effects of misconceptions and instructional method on transfer

Concerning transfer, the impact of misconceptions did not differ between learners who read the standard text and learners who read the refutation text. This result stands in contrast to the finding of a previous study that reading a refutation text compensates for the detrimental influence of misconceptions on transfer found with a standard text (Prinz et al. 2019). However, it needs to be considered that there was no significant negative effect of misconceptions on transfer in the group that read the standard text in the present study, presumably reflecting that it was a strong control group. Specifically, in addition to the standard text, the learners in this group were provided with a sheet that listed the targeted misconceptions so that they received the same amount of information as the learners in the other groups. This might have helped the learners who held some of these misconceptions to engage in certain effective learning processes as well. That is, they were to some extent able to use what they had read about the correct information concerning the misconceptions when it came to the application of this information. Thus, their transfer performance was quite similar to that of the learners in the control group who held fewer misconceptions. In turn, there was no differential impact of misconceptions on transfer in the control group versus the refutation-text group.

The impact of misconceptions on transfer differed, however, between learners who read the standard text and learners who completed the think sheet. As just described, misconceptions had no significant negative effect on transfer when learners read the standard text. Surprisingly, though, when learners completed the think sheet, those who had more misconceptions in the beginning achieved even better transfer performance than learners who had fewer misconceptions. Although incorrect prior knowledge is often regarded as detrimental for knowledge acquisition, it has been emphasized that such knowledge can sometimes also be productive for learning. In line with the constructivist perspective that new knowledge is built from prior knowledge, this is because misconceptions can function as important resources that learners can transform and refine into normatively correct conceptions. Hence, given appropriate instruction, misconceptions can serve as anchors in the process of building a correct understanding (e.g., Clark and Linn 2013; diSessa 2014; Smith et al. 1994). Apparently, think sheets represent such an instructional method that supports learners with misconceptions not only in acquiring an accurate comprehension but particularly also in setting this comprehension into practice. This finding also corroborates the utility of generative tasks for promoting a deep understanding that learners can

apply in new situations (see, e.g., Chi 2009; Fiorella and Mayer 2016). To conclude, when the aim of instruction is that learners revise their misconceptions and accomplish transfer of their knowledge, think sheets seem especially useful.

Effects of misconceptions and instructional method on bias of comprehension predictions

When reading the standard text, learners with more misconceptions overestimated their comprehension more strongly than learners with fewer misconceptions. In contrast, this effect was not present when learners read the refutation text or completed the think sheet. This result replicates the finding of a prior study that refutation texts can counteract the detrimental impact of misconceptions on the accuracy of learners' comprehension predictions (Prinz et al. 2019). In addition, this result extends prior research by showing that also think sheets are useful to prevent learners with misconceptions from overestimating their comprehension. This outcome indicates that the coactivation principle of the KReC framework (Kendeou and O'Brien 2014) not only represents an important condition for attaining more adequate comprehension but also more accurate metacomprehension. Specifically, both refutation texts and think sheets seem to facilitate coactivation and conceptual-change processes. In the context of these processes, learners might inevitably monitor and evaluate their understanding more carefully. This might in turn promote their use of judgment cues that tap their situation model instead of cues such as the pure amount of textual information retrievable from memory (e.g., Dunlosky et al. 2005). In fact, situation-model cues might become more salient with these tasks. Accordingly, it has been revealed that generative tasks, such as constructing concept maps and completing diagrams, help learners to predict their comprehension more accurately by providing situation-model cues, for example, how well they can depict relations reported in a text (see, e.g., Griffin et al. 2019; Prinz et al. 2020; Van de Pol et al. 2020). The generative task of completing a think sheet might also provide learners with cues that are indicative of their situation model, such as how well they can summarize the correct information from the text on the think sheet. Presumably, although reading a refutation text is not a generative task, because it requires active integration processes, it might likewise promote the use of situation-model cues.

Effects of misconceptions and instructional method on bias of transfer predictions

The influence of misconceptions on the accuracy of predictions of transfer did not differ between learners who read the standard text and learners who read the refutation text. This result contradicts the finding of a previous study that, compared with reading a standard text, reading a refutation text prevents learners with a higher number of misconceptions from overestimating their transfer (Prinz et al. 2019). Yet, in the present study, there was no significant detrimental impact of misconceptions on prediction accuracy of transfer in the group that read the standard text, ruling out the possibility to find a compensatory effect. However, the influence of misconceptions on the accuracy of predictions of transfer differed between learners who read the standard text and learners who completed the think sheet. As just stated, when learners read the standard

text, misconceptions had no significant effect. In contrast, when learners completed the think sheet, the more misconceptions they had in the beginning, the more strongly they underestimated their ability to transfer their acquired comprehension. Actually, in all three groups, the learners on average underestimated their transfer (see Table 2). In the experiment, the learners made the predictions of their comprehension and transfer performance simultaneously. Hence, it seems likely that, compared with comprehending a text, they considered the transfer of that comprehension to be more difficult, while overrating the difference in the difficulty between the two tasks. This has apparently especially been the case when learners who had more misconceptions completed the think sheet. Completing the think sheet made them aware of their misconceptions and required them to actively search for the correct information during reading, presumably even increasing the perceived difficulty of knowledge transfer. Consequently, although completing the think sheet actually supported them in transferring their acquired comprehension, they were unconfident in their ability to do so.

Effects of misconceptions and instructional method on response-confidence accuracy

With regard to response confidence, learners who completed the think sheet were more accurate than learners who read the standard or the refutation text, irrespective of how many misconceptions they had. This means that the think-sheet learners were generally more sensitive to the correctness of their knowledge than the standard- and refutation-text learners. This novel finding underlines the effectiveness of think sheets for promoting learners' metacognitive monitoring during testing. In fact, in the think-sheet group, the response-confidence accuracy was as high as .72, reflecting a substantial correlation between learners' judgments and performance. When studying, the think-sheet task required the learners to actively extract and compare the correct textual information with misconceptions. Hence, when judging the correctness of their responses to the test questions, the learners might also have mentally contrasted the answer they chose to the textual information. As a consequence, they more reliably realized when their answer could have been incorrect or was likely correct, resulting in greater response-confidence accuracy.

Delayed effects of misconceptions and instructional method on comprehension and transfer

At delayed testing, we did not find effects of misconceptions and instructional method on comprehension or transfer. This result seems to stand in contrast to previous findings that the beneficial effects of refutation texts and think sheets on learning are persistent (e.g., Dole and Smith 1989; Kendeou et al. 2014). However, this result reflects that the negative impact of misconceptions on learning was diminished in the control group at delayed testing. Presumably, the learners with misconceptions in this group particularly benefited from repeatedly completing test questions on the topic. In fact, research on the testing effect has shown that retrieval practice through answering test questions supports learning (see, e.g., Dunlosky et al. 2013; Roediger and Butler 2011). Additional explorative analyses revealed that, for all groups, rather than learners' misconceptions in the beginning, their knowledge acquired in the immediate session affected their performance in the delayed session. Moreover, comparing performance at immediate and delayed testing revealed no significant decline in comprehension or transfer, and this was the case independent of

condition or the number of misconceptions held in the beginning. Consequently, the beneficial influences of reading a refutation text and completing a think sheet appear to be quite stable.

Limitations and future directions

A few aspects have to be considered when interpreting the current findings. First, we assumed that, like refutation texts, think sheets promote coactivation and conceptual-change processes and the use of valid judgment cues. However, the actual mechanisms underlying the effects of completing a think sheet are unclear. Therefore, future research should apply process measures, such as the think-aloud methodology, eye-tracking, or self-reports. For example, participants could be asked to report their experiences during studying (e.g., their thoughts and emotions such as confusion) and the bases for their judgments (e.g., which cues they used). Second, there are several ways in which think sheets can be implemented in instruction. In the present study, the misconceptions about covariance were already entered on the think sheet to keep this aspect standardized. Another approach is that learners record their own particular prior beliefs concerning the questions on a think sheet (Dole and Smith 1989; see also Brummer and Macceca 2013; Fang et al. 2010). When learners activate and write down their misconceptions before being confronted with the correct information in a text, this might even be more beneficial for inducing coactivation and conceptual-change processes. A further possibility is that, after filling in the correct information from a text on the think sheet, learners are explicitly prompted to contrast their prior beliefs and the textual information. For example, in the study by Dole and Smith (1989), the think sheet contained additional columns that asked the students to contrast the two pieces of information. Third, in our study, the participants had to complete the think sheet while reading the text because this was supposed to effectively enhance their comprehension monitoring and help them become aware of knowledge deficits to be remedied (cf. Lachner et al. 2020). Yet, previous research has shown that some generative tasks, such as writing summaries, listing keywords, and completing diagrams, particularly promote metacomprehension accuracy when they are completed after a delay of a few minutes (see, e.g., Griffin et al. 2019; Prinz et al. 2020; Van de Pol et al. 2020). A learner's memory for details decays quite rapidly, whereas the situation model is more stable (Kintsch et al. 1990). Hence, when executed after a short delay, generative tasks can provide learners with enhanced access to cues related to their situation model with less interference from details. Thus, it would be interesting to investigate the impact of think sheets when learners complete them some time after reading. Fourth, some results were only marginally significant. However, corresponding effect sizes were of medium size, and post hoc power analyses showed that the statistical power for the analyses was below the common threshold of .80. The lacking power might explain why the effect sizes were not statistically significant at the .05 level. Future studies should replicate and extend the present findings by using a larger sample. Fifth, we decided to conduct the delayed session about 7 days after the immediate session. We chose this rather short time interval to avoid a large drop out due to the potential unavailability of participants (e.g., due to the semester break, internships). In addition, implementing a long delay covers the risk of missing effects in case they diminish. Nonetheless, to examine to what extent the effects of the instructional methods are stable over a longer period of time, future studies should deploy a lag of several weeks or months.

Implications and conclusion

The KReC framework (Kendeou and O'Brien 2014) suggests that the coactivation of a misconception with the correct information is a central prerequisite for conceptual change because only then can the correct information be further integrated and activated in a learner's knowledge structures. So far, this assumption has been investigated by using refutation texts (e.g., Ariasi and Mason 2011, 2014; Diakidoy et al. 2011; Kendeou et al. 2011; Kendeou and Van den Broek 2007; McCrudden 2012; McCrudden and Kendeou 2014; Prinz et al. 2019; Van den Broek and Kendeou 2008). Therefore, by using a novel, innovative instructional method, namely think sheets, to induce these processes, our study extends prior research and reinforces the importance of the coactivation principle of the KReC framework (Kendeou and O'Brien 2014). In addition, our study broadens the theoretical and empirical base by showing that coactivation is not only an important condition for improving comprehension but also metacomprehension accuracy when learners have misconceptions. Through the coactivation of a misconception and the correct information, learners are likely to monitor and evaluate their comprehension more carefully, which might facilitate the use of more valid judgment cues.

It has to be considered, though, that concerning the reduction of misconceptions also the control group was effective. The control group in this study received the same amount of information as the experimental groups. Specifically, they were provided not only with a standard text but also with a sheet that listed the targeted misconceptions. This might have led to some coactivation and conceptual-change processes as well, reducing the misconceptions. Hence, even the rather simple and easy-to-implement method of providing learners with a list of misconceptions before confronting them with a text on the correct information can be effective to debunk misconceptions and might be used in instruction for this purpose. However, for supporting learners who held misconceptions not only in reducing the misconceptions but also in acquiring the correct comprehension and accurately judging this comprehension, the methods of reading a refutation texts and completing a think sheet were more beneficial. Thus, these two methods have additional advantages that speak in favor of their usage in instruction. Moreover, this outcome indicates that think sheets represent an eligible instructional alternative to refutation texts. It seems reasonable to alternate the application of the methods in instruction to keep up learners' engagement (see, e.g., Smith and Ragan 2005). Yet, think sheets apparently have some further benefits. Specifically, completing the think sheet facilitated the knowledge transfer of learners who had misconceptions and generally supported learners in accurately discriminating between correct and incorrect responses. Hence, when the goal of instruction is that learners can transfer their knowledge or that they become aware of knowledge deficits during practice testing, think sheets might be preferred. In sum, the present study reinforces the importance of directly addressing misconceptions in instruction to support effective learning and emphasizes the utility of two instructional methods for this purpose, namely refutation texts and think sheets.

Funding Open Access funding enabled and organized by Projekt DEAL. This study was funded by the Wissenschaftliche Gesellschaft Freiburg (Scientific Society Freiburg).

Data availability The data and materials are available from the corresponding author on request.

Code availability Not applicable.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The study was implemented in alignment with the principles of ethical and professional conduct of the German Psychological Society (DGPs). An ethical approval was not mandatory for the type of collected data at the time when the study was conducted (December 2017–February 2018).

Consent to participate The data was collected and saved anonymously. Informed consent was obtained from all participants.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

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(5), 581–601. <https://doi.org/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(3), 493–523. <https://doi.org/10.1007/s10763-013-9494-9>
- Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior*, 52, 293–306. <https://doi.org/10.1016/j.chb.2015.05.038>
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612–637. <https://doi.org/10.1037/0033-2909.128.4.612>
- Batanero, C., Estepa, A., & Godino, J. D. (1996). Evolution of students' understanding of statistical association in a computer-based teaching environment. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 191–205). Voorburg, the Netherlands: International Statistical Institute.
- Beker, K., Kim, J., Van Boekel, M., Van den Broek, P., & Kendeou, P. (2019). Refutation texts enhance spontaneous transfer of knowledge. *Contemporary Educational Psychology*, 56, 67–78. <https://doi.org/10.1016/j.cedpsych.2018.11.004>
- Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, 17(5), 564–577. <https://doi.org/10.1016/j.learninstruc.2007.09.007>
- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* [Statistics for human and social scientists] (7th ed.). Berlin, Germany: Springer.
- 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), 11–30. <https://doi.org/10.1037/a0032627>
- Brummer, T., & Maccetta, S. (2013). *Reading strategies for mathematics*. Huntington Beach, CA: Shell Education.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition & Instruction*, 8(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2
- Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105. <https://doi.org/10.1111/j.1756-8765.2008.01005.x>

- Clark, D. B., & Linn, M. C. (2013). The knowledge integration perspective: Connections across research and education. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (2nd ed., pp. 520–538). New York, NY: Routledge.
- 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(2), 164–174. <https://doi.org/10.1016/j.cedpsych.2014.03.006>
- De Bruin, A. B. H., Thiede, K. W., Camp, G., & Redford, J. (2011). Generating keywords improves meta-comprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*, 109(3), 294–310. <https://doi.org/10.1016/j.jecp.2011.02.005>
- 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. [https://doi.org/10.1016/S0361-476X\(02\)00039-5](https://doi.org/10.1016/S0361-476X(02)00039-5)
- Diakidoy, I.-A.N., Mouskounti, T., & Ioannides, C. (2011). Comprehension and learning from refutation and expository texts. *Reading Research Quarterly*, 46(1), 22–38. <https://doi.org/10.1598/RRQ.46.1.2>
- diSessa, A. (2014). The construction of causal schemes: Learning mechanisms at the knowledge level. *Cognitive Science*, 38(5), 795–850. <https://doi.org/10.1111/cogs.12131>
- Dole, J. A., & Smith, E. L. (1989). Prior knowledge and learning from science text: An instructional study. In S. McCormick & J. Zutell (Eds.), *Cognitive and social perspectives for literacy research and instruction. Thirty-eighth yearbook of the National Reading Conference* (pp. 345–352). National Reading Conference, Chicago, IL. <https://files.eric.ed.gov/fulltext/ED313664.pdf>
- Donovan, A. M., Zhan, J., & Rapp, D. N. (2018). Supporting historical understandings with refutation texts. *Contemporary Educational Psychology*, 54, 1–11. <https://doi.org/10.1016/j.cedpsych.2018.04.002>
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4–58. <https://doi.org/10.1177/1529100612453266>
- 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(4), 551–565. <https://doi.org/10.1016/j.jml.2005.01.011>
- Fang, Z., Lamme, L. L., & Pringle, R. M. (2010). *Language and literacy in inquiry-based science classrooms, grades 3–8*. Thousand Oaks, CA: Corwin.
- Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, 28(4), 717–741. <https://doi.org/10.1007/s10648-015-9348-9>
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221–233. <https://doi.org/10.1037/h0057532>
- Griffin, T. D., Jee, B. D., & Wiley, J. (2009). The effects of domain knowledge on metacomprehension accuracy. *Memory & Cognition*, 37(7), 1001–1013. <https://doi.org/10.3758/MC.37.7.1001>
- Griffin, T. D., Wiley, J., & Salas, C. R. (2013). Supporting effective self-regulated learning: The critical role of monitoring. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (Vol. 28, pp. 19–34). New York, NY: Springer.
- Griffin, T. D., Mielicki, M. K., & Wiley, J. (2019). Improving students' metacomprehension accuracy. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (1st ed., pp. 619–646). New York, NY: Cambridge University Press.
- Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, 28(2), 116–159. <https://doi.org/10.2307/747886>
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). New York, NY: Guilford Press.
- Huck, S. W. (2016). *Statistical misconceptions*. New York, NY: Routledge.
- Kramarski, B., & Friedman, S. (2014). Solicited versus unsolicited metacognitive prompts for fostering mathematical problem solving using multimedia. *Journal of Educational Computing Research*, 50(3), 285–314. <https://doi.org/10.2190/EC.50.3.a>
- Kendeou, P., Braasch, J. L. G., & Bråten, I. (2016). Optimizing conditions for learning: Situating refutations in epistemic cognition. *The Journal of Experimental Education*, 84(2), 245–263. <https://doi.org/10.1080/00220973.2015.1027806>
- Kendeou, P., Muis, K. R., & Fulton, S. (2011). Reader and text factors in reading comprehension processes. *Journal of Research in Reading*, 34(4), 365–383. <https://doi.org/10.1111/j.1467-9817.2010.01436.x>
- 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 science* (pp. 353–377). Cambridge, MA: MIT University Press.
- Kendeou, P., & Van den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology*, 97(2), 235–245. <https://doi.org/10.1037/0022-0663.97.2.235>
- Kendeou, P., & Van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, 35(7), 1567–1577. <https://doi.org/10.3758/BF03193491>
- Kendeou, P., Walsh, E. K., Smith, E. R., & O'Brien, E. J. (2014). Knowledge revision processes in refutation texts. *Discourse Processes*, 51(5–6), 374–397. <https://doi.org/10.1080/0163853X.2014.913961>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.
- Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, 29(2), 133–159. [https://doi.org/10.1016/0749-596X\(90\)90069-C](https://doi.org/10.1016/0749-596X(90)90069-C)
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126(4), 349–370. <https://doi.org/10.1037/0096-3445.126.4.349>
- Kowalski, P., & Taylor, A. K. (2017). Reducing students' misconceptions with refutational teaching: For long-term retention, comprehension matters. *Scholarship of Teaching and Learning in Psychology*, 3(2), 90–100. <https://doi.org/10.1037/stl0000082>
- Lachner, A., Backfisch, I., Hoogerheide, V., Van Gog, T., & Renkl, A. (2020). Timing matters! Explaining between study phases enhances students' learning. *Journal of Educational Psychology*, 112(4), 841–853. <https://doi.org/10.1037/edu0000396>
- Lassonde, K. A., Kendeou, P., & O'Brien, E. J. (2016). Refutation texts: Overcoming psychology misconceptions that are resistant to change. *Scholarship of Teaching and Learning in Psychology*, 2(1), 62–74. <https://doi.org/10.1037/stl0000054>
- Liu, T.-C., Lin, Y.-C., & Tsai, C.-C. (2009). Identifying senior high school students' misconceptions about statistical correlation, and their possible causes: An exploratory study using concept mapping with interviews. *International Journal of Science and Mathematics Education*, 7(4), 791–820. <https://doi.org/10.1007/s10763-008-9142-y>
- 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. <https://doi.org/10.1037/0022-0663.100.2.291>
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.
- McCrudden, M. T. (2012). Readers' use of online discrepancy resolution strategies. *Discourse Processes*, 49(2), 107–136. <https://doi.org/10.1080/0163853X.2011.647618>
- McCrudden, M. T., & Kendeou, P. (2014). Exploring the link between cognitive processes and learning from refutational text. *Journal of Research in Reading*, 37(1), 116–S140. <https://doi.org/10.1111/j.1467-9817.2011.01527.x>
- McNamara, D. S., & Magliano, J. P. (2009). Toward a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 297–384). San Diego, CA: Elsevier.
- Mikkilä-Erdmann, M. (2001). Improving conceptual change concerning photosynthesis through text design. *Learning and Instruction*, 11(3), 241–257. [https://doi.org/10.1016/S0959-4752\(00\)00041-4](https://doi.org/10.1016/S0959-4752(00)00041-4)
- Moritz, J. (2004). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227–255). Dordrecht, the Netherlands: Springer.
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). New York, NY: Academic Press.
- Nussbaum, E. M., Cordova, J. R., & Rehmat, A. P. (2017). Refutation texts for effective climate change education. *Journal of Geoscience Education*, 65(1), 23–34. <https://doi.org/10.5408/15-109.1>
- Organisation for Economic Co-operation and Development. (2006). *PISA 2006—Schulleistungen im internationalen Vergleich: Naturwissenschaftliche Kompetenzen für die Welt von morgen* [PISA 2006: Science competencies for tomorrow's world]. Retrieved from <http://www.oecd.org/pisa/39728657.pdf>.
- Prinz, A., Golke, S., & Wittwer, J. (2018). The double curse of misconceptions: Misconceptions impair not only text comprehension but also metacomprehension in the domain of statistics. *Instructional Science*, 46(5), 723–765. <https://doi.org/10.1007/s11251-018-9452-6>

- Prinz, A., Golke, S., & Wittwer, J. (2019). Refutation texts compensate for detrimental effects of misconceptions on comprehension and metacomprehension accuracy and support transfer. *Journal of Educational Psychology*, 111(6), 957–981. <https://doi.org/10.1037/edu0000329>
- Prinz, A., Golke, S., & Wittwer, J. (2020). To what extent do situation-model-approach interventions improve relative metacomprehension accuracy? Meta-analytic insights. *Educational Psychology Review*, 32(4), 917–949. <https://doi.org/10.1007/s10648-020-09558-6>
- Rawson, K. A., O'Neil, R., & Dunlosky, J. (2011). Accurate monitoring leads to effective control and greater learning of patient education materials. *Journal of Experimental Psychology: Applied*, 17(3), 288–302. <https://doi.org/10.1037/a0024749>
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20–27. <https://doi.org/10.1016/j.tics.2010.09.003>
- Schleinschok, K., Eitel, A., & Scheiter, K. (2017). Do drawing tasks improve monitoring and control during learning from text? *Learning and Instruction*, 51, 10–25. <https://doi.org/10.1016/j.learninstruc.2017.02.002>
- Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning*, 4(1), 33–45. <https://doi.org/10.1007/s11409-008-9031-3>
- Smith, P. L., & Ragan, T. J. (2005). *Instructional design* (3rd ed.). Hoboken, NJ: Wiley.
- Smith, J. P., III., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, 3(2), 115–163. https://doi.org/10.1207/s15327809jls0302_1
- 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. <https://doi.org/10.1037/0022-0663.95.1.66>
- Thiede, K. W., Redford, J. S., Wiley, J., & Griffin, T. D. (2012). Elementary school experience with comprehension testing may influence metacomprehension accuracy among seventh and eighth graders. *Journal of Educational Psychology*, 104(3), 554–564. <https://doi.org/10.1037/a0028660>
- 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. <https://doi.org/10.1007/s10763-010-9203-x>
- Trevors, G. J., Kendeou, P., Bråten, I., & Braasch, J. L. G. (2017). Adolescents' epistemic profiles in the service of knowledge revision. *Contemporary Educational Psychology*, 49, 107–120. <https://doi.org/10.1016/j.cedpsych.2017.01.005>
- Van Boekel, M., Lassoend, K. A., O'Brien, E. J., & Kendeou, P. (2017). Source credibility and the processing of refutation texts. *Memory & Cognition*, 45(1), 168–181. <https://doi.org/10.3758/s13421-016-0649-0>
- 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. <https://doi.org/10.1002/acp.1418>
- Van de Pol, J., Van Loon, M. H., Van Gog, T., Braumann, S., & De Bruin, A. B. H. (2020). Mapping and drawing to improve students' and teachers' monitoring and regulation of students' learning from text: Current findings and future directions. *Educational Psychology Review*, 32(4), 951–977. <https://doi.org/10.1007/s10648-020-09560-y>
- Weingartner, K. M., & Masnick, A. M. (2019). Refutation texts: Implying the refutation of a scientific misconception can facilitate knowledge revision. *Contemporary Educational Psychology*, 58, 138–148. <https://doi.org/10.1016/j.cedpsych.2019.03.004>
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Erlbaum.
- Witrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist*, 24(4), 345–376. https://doi.org/10.1207/s15326985ep2404_2
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612–637. <https://doi.org/10.1037/0033-2909.128.4.612>