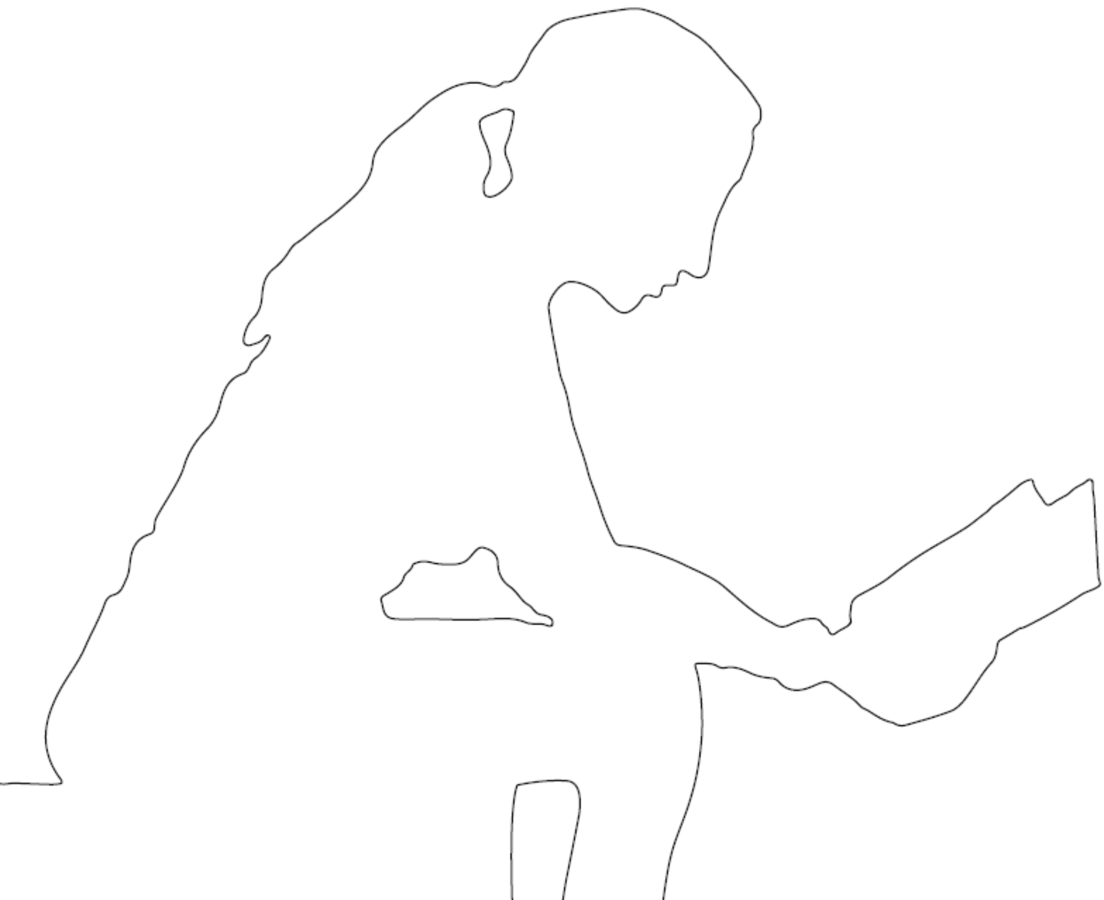
Author: Swart, E.K.
Title: Mind the reading mind: a multifaceted and methodologically diverse approach to investigating the role of attentional control and feedback in reading comprehension
Issue Date: 2021-06-08



Chapter 4

Supporting Learning From Text: A Meta-Analysis on the Timing and Content of Effective Feedback

Published as:

Swart, E. K., Nielen, T. M. J., & Sikkema-de Jong, M. T. (2019). Supporting learning from text: A meta-analysis on the timing and content of effective feedback. *Educational Research Review*, *28,* 100296. doi:10.1016/j.edurev.2019.100296

Abstract

The aim of the present meta-analysis was to examine the effects of feedback on learning from text in conventional readers (ranging from primary school students to university students). Combining 104 contrasts of conditions of reading texts with and without feedback, including 6,124 participants, using the random effects model resulted in a positive effect of feedback on learning from text ($g+ = 0.35$). Moderator analyses showed that feedback is particularly effective if provided directly after reading, but less so when provided during reading. If feedback is provided directly after reading, elaborate feedback and knowledge-of-correct-response feedback were more effective than knowledge-of-response feedback. If feedback is provided during reading, no differences are found between the effects of different types of feedback. Additionally, computer-delivered feedback is more beneficial for learning from text than non-computer-delivered feedback. Implications for optimizing conditions to support learning from text are discussed.

Keywords: learning from text, reading comprehension, feedback, computer-assisted learning, cognitive load theory

Introduction

Although reading comprehension is an essential skill for learning from texts, international reading research consistently shows that on average 1 out of 5 students at the age 15 has not reached the basic level of reading comprehension that is required to effectively learn from texts and to function well in society (OECD, 2010, 2014, 2016). This means that these students are not sufficiently able to extract the main idea from a text, to understand relations between parts of texts and/or to link information in a text to background knowledge or personal experiences (OECD, 2016). That is, these students are not able to learn from a text because they do not succeed to create a complete and coherent mental model of the text they are reading (Kintsch, 1986). The alarming number of struggling readers and the fact that we have not been able to reduce this number in the last decade emphasizes the need to invest in the development of effective teaching strategies that support learning from text, not only for young children but also for adolescent readers (see Edmonds et al., 2009).

Effective teaching strategies for developing the reading comprehension skills that are essential for learning from text should contain three vital elements: (a) sustained deliberate practice (i.e., purposeful and repeated practice with texts in varying domains and genres), (b) strategy instruction (i.e., guidance and information on how to select and use appropriate and effective reading strategies during reading), and (c) individual feedback (i.e., tailored to the needs of the individual reader; Crossley & McNamara, 2017). There is a large base of literature on effective reading programs that are aimed at teaching readers strategies in general for better reading comprehension (for reviews, see e.g., Dole et al., 1991; Fuchs & Fuchs, 2005; Paris & Jacobs, 1984; Slavin et al., 2008, 2009). Another line of research in the scientific literature on reading comprehension is focused on supporting learning from text 'on the job'. That is, students are for example given comprehension questions during or directly after reading a text and receive individual feedback on their answers in order to facilitate learning from text (see e.g., Kulhavy, 1977). Providing students constantly with individual feedback is a time consuming activity for which teachers often do not have enough time. Information and computer technologies might provide possibilities to deliver personalized instruction and feedback to children and students while reading to optimize learning from text (e.g., Nielen et al., 2017).

Decades of research have shown the positive effects of feedback on learning performance (e.g., Azevedo & Bernard, 1995; Hattie 2012; Jaehnig & Miller, 2007; Kluger & DeNisi, 1996; Kulhavy, 1977; Shute, 2008; van der Kleij et al., 2012, 2015). Additionally, based on more than 900 meta-analyses on schooling effects, Hattie (2012) concluded that feedback is one of the top-10 positive influences on learning, even though studies on the effects of feedback show very much variability. This variability not only concerns the magnitude of the reported effects of feedback among studies, but also the ways in which feedback is provided to students. In line with these review studies, we consider feedback as individualized information that is communicated to the reader in response to the reader's performance (e.g., answers on comprehension questions or words that have to be filled in while reading a text) with the intention to enhance learning. It is important to get a better understanding of how providing feedback might help or obstruct learning from text because most learning in schools is still the result of understanding the information that is presented in written form.

According to the *Feedback Intervention Theory* of Kluger and DeNisi (1996), the effect of feedback may depend on the task or skill targeted by the feedback, the design of the feedback, and other situational variables (i.e., personality or methodological variables). Contrary to previous review studies in which the effects of feedback on a broad range of targeted tasks and skills were investigated qualitatively (e.g., Jaehnig & Miller, 2007) and quantitatively (e.g., Kluger & DeNisi, 1996; Shute, 2008), the present quantitative meta-analysis has a narrower scope, investigating the effects of feedback specifically targeted at learning from text. In the included studies the feedback was given during or after the reading task. Studies on reading programs of which feedback was only an integral part, like reciprocal teaching, tutorial dialogues or reading with the help of (voluntary) reading tutors were outside the scope of the present meta-analysis (e.g., see Palinscar & Brown, 1984). To better understand when and why feedback is or is not effective to support learning from text we tested the impact of three feedback design variables: (a) timing of the feedback, (b) type of feedback and (c) the means of providing feedback to the reader. The third type of variables mentioned by Kluger and DeNisi (1996), situational variables (i.e., personality variables and methodological variables), are taken into account when checking the results of the meta-analysis for possible biases (e.g., age group of the participants, instructional design of the texts that were read in the studies, random assignment of participants to conditions, experimental design of the studies).

Timing of Feedback

Although feedback appears to be a vital element that supports learning from text, the scientific literature differs on the optimal timing for providing feedback. In previous studies the focus regarding the issue of timing of feedback has mainly been on the comparison of immediate feedback (i.e., during or right after a learning task) versus delayed feedback (i.e., feedback provided some hours or days after the learning task; Dempsey & Wager, 1988; Hattie & Timperley, 2007; Metcalfe et al., 2009; Mory, 2004; Shute, 2008). In the present study we take a new approach to look at differences in timing of feedback, namely by comparing two forms of immediate feedback: feedback that is provided on student responses to questions or tasks during reading (i.e., after reading segments of a text) and feedback that is provided on student responses to questions or tasks directly after reading (i.e., after finishing a whole text). Some studies show positive effects on learning from text if feedback is provided during reading (e.g., Llorens et al., 2014) whereas other studies show no effect (e.g., Farragher & Yore, 1977) or detrimental effects of feedback on learning from text (e.g., Fernald & Jordon, 1991). Similarly, results are mixed among studies in which the effects of feedback directly after reading on learning from text are tested (e.g., Adams & Strickland, 2012; Butler et al., 2013; Saunders, 1998).

From a theoretical point of view, one could argue for both: providing feedback during reading and/or providing feedback directly after reading to support learning from text. Feedback could best be provided during reading based on the idea that feedback helps readers to evaluate and adjust their knowledge during reading, so that flaws in understanding (i.e., an inadequate mental representation of the text) can be detected and corrected as soon as possible. From this perspective, feedback can act as scaffold for comprehension monitoring. This metacognitive skill stimulates the reader to take corrective actions in cases of flaws in understanding, which is an essential skill for learning from text (Baker & Brown, 1984). The preference for feedback during reading corresponds with the five-stage model by Bangert-Drowns et al. (1991) describing the state of a learner when receiving feedback. In the first three stages of the model the reader (1) activates knowledge, interests, goals and/or beliefs about self-efficacy, (2) applies search and retrieval strategies that are activated by a question or task, and (3) gives a response to a question or task. According to the model, feedback follows after the response of the reader (i.e., after the third stage) and triggers the next two stages in which the reader (4) evaluates and (5) adjusts task-relevant knowledge, interests, goals or beliefs about self-efficacy. In line herewith, Mullet and Marsh (2016) formulated two guidelines which feedback has to comply to in order to effectively enhance learning: it needs to help learners (a) to notice errors and (b) to correct them as soon as possible. According to the five-stage model of Bangert-Drowns et al. (1991) and the guidelines of Mullet and Marsh (2016) feedback should thus function as a 'forget-cue' to prevent errors from persisting in memory, meaning that feedback during reading would be more effective to support learning from text as compared to feedback after reading.

From a cognitive load perspective, it is important to design feedback in a way that reduces the reader's cognitive load, in other words the information that has to be held in working memory during a task, while reading a text to enhance learning (Sweller, 1994; Sweller et al., 1998). From this perspective, one could argue for both feedback during reading and/or directly after reading. On the one hand, interrupting reading with tasks (e.g., comprehension questions or blanks in a text that have to be filled in) and subsequent feedback during reading increases the extraneous cognitive load of a reading task compared to reading a text when questions and feedback are provided directly after reading. Providing feedback during reading forces the reader to switch attention between the building of a coherent mental model of a text and the processing of the question/task and the subsequent feedback. In other words, providing feedback during reading requires the reader to multitask by switching between reading the text and executing the tasks or answering questions and processing the provided feedback (see Subrahmanyam et al., 2013). In line with this reasoning, it might be better to provide feedback directly after reading instead of during reading, so that students' limited working memory capacity can fully be deployed to create a mental model of the text (see also Kluger & DeNisi, 1996). On the other hand, questions and feedback during reading may help the reader to construct schemas of the information in the text (i.e., connecting parts of information within the text with each other or with background knowledge). By schematizing the information in the text, several pieces of information from the text can be chunked as a single element. In line with this idea, questions and feedback during reading help readers to construct these chunks of information from the text, which reduces cognitive load, and integrate them in their mental model of the text (see Paas et al., 2004).

Only one study (Peverly & Wood, 2001) on the effects of feedback on learning from text directly compared the effects of feedback during reading and feedback directly after reading within one experiment. Students from grade 9 and 11 in two feedback conditions (i.e., during and directly after reading) significantly outperformed students in the control conditions on an unstandardized reading comprehension post-test that assessed learning from texts of similar length and readability level as the stories in the intervention. Additionally, students receiving feedback during reading outperformed students who answered questions and received feedback directly after reading. No differences in performance were found on a standardized reading comprehension post-test between the feedback conditions versus the control group nor between the two feedback conditions. However, this was a small-scale study with only ten students in each condition and further research that directly compares the effects of feedback during and after reading is needed.

Different Types of Feedback

Another aspect of providing feedback that is still under debate in the scientific literature on the effects of feedback on learning from text is the issue regarding the type of feedback that would enhance learning best (see e.g., Shute, 2008). Kulhavy (1977) was one of the first researchers who reviewed studies aimed at testing the effects of feedback on learning from text. In his review, Kulhavy described that the composition of feedback can range along a "continuum from the simplest *Yes-No* format to the presentation of substantial corrective or remedial information that may extend the response content, or even add new material to it" (Kulhavy, 1977, p. 212). More than a decade later, Kulhavy and Stock (1989) argued that when studying the effects of feedback on learning from an information processing perspective, the content of a feedback message can be subdivided into two parts: *verification* and *elaboration*. Verification consists of a simple yes/no or right/wrong statement. In most studies, feedback that only includes verification is called *Knowledge-of-Response* (*KOR*; for a review see Shute, 2008). All information added to the verification statement can be seen as the elaboration part of the feedback message. According to Kulhavy and Stock (1989), the elaboration part of the feedback can include three types of information with increasing complexity: task-specific information (i.e., the correct answer), instruction-based information (e.g., explaining why an answer is correct or repeating the part of a text that contained information regarding the correct answer) and extra-instructional information (i.e., providing information, examples or analogies that were not present in the original learning material) respectively. When the verification part of the feedback is elaborated with the correct answer, the feedback message is often referred to as *Knowledge-of-Correct-Response* (*KCR*). Feedback that additionally includes instruction-based or extra-instructional information is often referred to as *Elaborated Feedback (EF)*. Kluger and DeNisi (1996) referred to the different amounts of information included in a feedback message as feedback specificity. More than a decade after the first review of studies on the effects of feedback by Kulhavy (1977), Kulhavy and Stock (1989) still had to conclude that studies on the effects of different types of feedback on learning from text showed inconsistent results. Also, researchers of more recent review studies on the effects of feedback on learning came to the same conclusion (e.g., Jaehnig & Miller, 2007; Kluger & DeNisi, 1996; Mory, 2004; Shute, 2008; van der Kleij et al., 2015).

The inconsistent results on the effects of different types of feedback on learning from text found in the existing literature may relate to differences in the focus of the provided feedback. Hattie and Timperley (2007) argued for instance that feedback can work at four different levels: task level (focused on how well a task is performed), process level (focused on the strategies that are needed to perform a task), self-regulation level (e.g., focused on self-evaluation or comprehension monitoring), and self-level (usually praise; see also Hattie, 2012). Hattie and Timperley (2007) described KOR and KCR as being task level feedback, that would be effective for building surface knowledge, whereas elaborated feedback (EF), which includes information on the process or self-regulation level, would be especially effective in enhancing deep processing of information. In line with this reasoning, van der Kleij et al. (2012) argued in a meta-analysis on the effects of feedback on learning in computer-based environments that providing EF was particularly effective for promoting higher-order information processing that requires not only literal recall of information, but also making inferences or transfer of knowledge. Based on the idea that learning from text is a form of higher-order information processing that goes beyond building surface knowledge such as storing facts (e.g., Craik & Lockhart, 1972), one would expect EF, compared to KOR or KCR, to be especially effective in enhancing learning from text. In contrast to this expectation however, in many recent studies in which the effects of different types of feedback were compared no differences were found in the effects of more or less elaborate feedback on learning from text (Golke et al. 2015; Llorens et al. 2014). Butler et al. (2013) further studied this unexpected lack of effect of EF by including different assessment levels for testing the effect of EF versus KCR. They showed that, compared to KCR, EF better enabled the students to answer inference questions. On a comprehension post-test assessing more superficial knowledge of the texts (i.e., definition questions) EF was not more effective than KCR in supporting learning from text. We hypothesize therefore that EF is especially effective when deep processing is required and that KCR is as effective for information processing on a surface level (see Graesser et al., 1997).

Means of Providing Feedback

Next to the timing and type of feedback, studies on the effects of feedback on learning from text also differ in the means of providing feedback to the reader. The emergence of computer-applications for reading instruction creates a wealth of possibilities to provide computer-delivered feedback during or directly after reading. Feedback can not only be provided using multiple modalities (e.g., text, audio, visuals or a tutor on screen), but also in a spatially and temporally integrated format. In other words, after answering a question or completing a task, the reader is only a mouse-click away from the feedback appearing on the screen. However, also before the computer era researchers have studied the effects of feedback on learning from text. In these early studies (e.g., Farragher & Yore, 1997; Feldhusen & Birt, 1962) feedback was non-computer-delivered, requiring participants to actively uncover the feedback messages themselves (for example removing a sticker that covers the feedback). The effects of non-computer-delivered feedback on learning from text found in these early studies are mixed. For example, in a study of Feldhusen and Birt (1962) college students read texts about teaching machines and programming language and had to fill in words in blanks in these texts during reading. Students had to check each answer by comparing their response to the answer on a feedback flap that was stored away in a separate folder. Students learned more from the text when they had to look at the feedback after each answer compared to a condition in which they had to fill in blanks during reading, but were not provided with feedback. Farragher and Yore (1997) performed a study in which 9th grade students read a text on environmental science and were asked a question after every 20 sentences. Students had to check their answers by using latent image pens that uncovered feedback responses that were printed with invisible ink. However, no effects of feedback were found on learning from the text. More recently, Clariana and Koul (2006) used a comparable approach; 10th and 11th grade students read expository science texts and had to answer questions during reading. After answering a question, students had to scratch a blue dot (like scratch-off lottery tickets) to make the correct answer visible. Students in feedback conditions outperformed students in a no feedback-condition on a comprehension post-test that assessed how much students had learned from the texts.

In a study of Lasoff (1981) a comparison was made between non-computer-delivered feedback and computer-delivered feedback. University students read five texts about programming language and answered multiple choice questions during reading. One group of students read the texts, answered questions and received feedback on separate sheets of paper and one group of students read the texts and answered the questions on the computer and also received computer-delivered feedback on their answers. The feedback in both conditions included the correct answer and explanations of why an answer was correct or hints when the answer was incorrect, so that the student could try again. No differences in learning were found between the feedback and no feedback conditions and no difference was found between the group that was provided with non-computer-delivered feedback and the group that was provided with computer-delivered feedback. However, this was a small-scale study with only ten to twelve students in each condition and more research is needed in which the effects of non-computer-delivered feedback and computer-delivered feedback is systematically compared.

Although there is a wide range of means of providing feedback, the cognitive load theory may help to predict which means will be most effective. That is, apart from the complexity of the text, extraneous factors such as the means of providing feedback are of influence on the overall cognitive load of the reading task (Sweller et al., 1998). According to the split-attention effect, working memory is challenged less when information is presented in a spatially and temporally integrated rather than in a separated format (e.g., Florax & Ploetzner, 2010). In line with this perspective, we hypothesize that computer-delivered feedback can more easily meet these requirements compared to non-computer-delivered and is, therefore, more effective in supporting learning from text.

Present Study

The aim of the present meta-analysis is to examine the effects of feedback on learning from text in conventional readers. In line with previous meta-analyses on the effects of feedback on learning (e.g., Azevedo & Bernard, 1995; Hattie 2012; Jaehnig & Miller, 2007; Kluger & DeNisi, 1996; Kulhavy, 1977; Shute, 2008; van der Kleij et al., 2015), we expected to find a positive effect of feedback on learning from text. In order to explain differences in the effects of feedback in previous studies, we investigated three characteristics of feedback related to design of the feedback: (1) timing of feedback, (2) types of feedback, and (3) the means by which feedback is provided to the reader. Regarding the effects of the timing of feedback, in line with the idea that feedback during reading helps the reader to chunk information (Paas et al., 2004) one could expect feedback during reading to be more effective. On the other hand, if feedback during reading appears to increase the cognitive load, because the reader has to multitask between reading and understanding the text and answering questions and processing feedback, than feedback that is provided directly after reading could be more effective to support learning from text (Subrahmanyam et al., 2013). Regarding the different types of the feedback, we expect more extensive forms of feedback, like elaborated feedback (EF) to be more effective than simple forms of feedback, like knowledge-of-response feedback (KOR), to support learning from text. We expected EF to be particularly effective when deep information processing is required for answering the questions about a text. Additionally, because it takes more effort and time, and therefore working memory capacity, to process non-computer-delivered feedback we expect, in line with the split attention hypothesis, computer-delivered feedback to be more effective in supporting learning from text than non-computer-delivered feedback.

Method

Inclusion and Exclusion Criteria

In line with previous review studies on the effects of feedback on learning (e.g., Kluger & DeNisi, 1996; Shute 2008; van der Kleij et al., 2015), we defined feedback as item-based individualized information that is communicated in *reaction* to the readers' response with the intention to enhance learning from the text. In the present study we looked at feedback that was provided during reading or directly after reading a text, concerning the reader's performance (i.e., answer to a question). Interactive options in a text that did not include information that reflects the individual performance of the reader, like providing pronunciation or glossary options were not considered as feedback (e.g., Davidson & Noyes, 1995). Additionally, feedback had to be systematically provided to the reader as opposed to incidentally whereby the methods of delivering feedback may differ (e.g., computer-delivered feedback, see Llorens et al, 2014; answers and/or hints masked with chemical ink, a flap or sticker, see Farragher & Yore, 1997 and Feldhusen & Birt, 1962; or separate answer sheets inserted in workbooks, see Clariana & Koul, 2006).

Studies on the effects of interventions of which feedback was only an integral part, like reciprocal teaching, tutorial dialogues or reading with the help of (voluntary) reading tutors were excluded from the meta-analysis. Studies were also excluded if feedback was already visible before answering questions or completing tasks, because in these cases feedback was not given in reaction of the reader's response (e.g., the copying condition in Anderson et al., 1972), if participants had to read in pairs or small groups and received the feedback per group of participants instead of individual feedback (e.g., MacGregor, 1988), or if participants had the option to turn off the feedback (e.g., Mikulecky, 1987).

In addition to the operational definition of feedback described above, studies were included in the meta-analysis if the following criteria were met:

(a) An intervention study had to be described, either using a between-subjects or within-subjects design, and a comparison had to be made between a feedback condition and a comparison condition.
(b) In the feedback condition, participants received feedback during reading or directly after reading a text.
(c) Participants in the control condition had to read a similar or comparable text, but without receiving feedback.
(d) The reading task had to concern reading narrative or informative texts, not lists of words or separate sentences. The texts could be accompanied by pictures or video-content.
(e) The text had to be read by participants themselves (requiring them to be conventional readers), not read to them by a computer or adult.
(f) At least one outcome measure of learning from text had to be reported. Learning from text could be measured using comprehension questions, recall or retelling tasks or tasks targeted at ordering text elements or pictures.
(g) Participants had no mental, physical, or sensory handicaps and were conventional readers.
(h) Reports of the studies had to be written in English.
(i) A sufficient amount of statistical information had to be reported to calculate effect sizes for the outcome measures that are relevant for the present meta-analysis.

Outcome measures for learning from text were included in the present meta-analysis if learning of information from the text(s) read during the intervention was assessed using comprehension questions (e.g., criterion test used in Anderson et al., 1972) or recall tasks (e.g., cued recall test used in Kealy & Ritzhaupt, 1962). Also, outcome measures for learning from new texts that were read as a part of a reading comprehension post-test were included (e.g., accuracy in answering comprehension questions about a text read in the final phase, which was different from the ones read in the training phase, in Llorens et al., 2014). Both measures developed by the researchers (e.g., achievement test used in Lasoff, 1981; comprehension questions used in Llorens et al, 2014) as well as standardized reading comprehension tests in which participants had to answer comprehension questions about texts they had to read during the test (e.g., Expository Text Comprehension Test and the Narrative Text Comprehension Test used in Sung et al., 2008) were included in the meta-analysis.

No restrictions were set for the inclusion of studies in the meta-analysis regarding the participant's age, country or origin, publication status or publication date of the reports of the studies. However, studies were excluded if the described samples or data overlapped with those of other studies in the meta-analysis (e.g., Wijekumar et al., 2013).

Search Strategies

To obtain eligible studies seven databases (PsycInfo, PsycArticles, ERIC, Proquest Dissertations and Theses Global, Web of Science, Linguistic and Language Behavior Abstracts and Google Scholar) were searched online up to March 2017 for journal articles, reports, conference proceedings, books, book chapters and dissertations describing studies investigating the effects of feedback on learning from text. Several combinations of search terms were used referring to books, e-books, literacy and reading on the one hand, and feedback, scaffolding, interactivity and tutoring on the other hand (see Appendix A for the exact search terms). Additionally, the reference lists of relevant handbooks on literacy, reading and technology (see Appendix B for an overview of searched handbooks) and review papers that came up in the online search were checked for studies to include in the meta-analysis. Subsequently, publication records of authors of papers that were found eligible for the meta-analysis were checked using Google Scholar. Over 15,000 (duplicates excluded) references and abstracts were checked for eligibility by the first and second author, of which 419 reports were checked in full-text. The described search strategies resulted in 60 eligible reports. For an overview of the search steps and the corresponding numbers of checked articles, see the PRISMA diagram in Appendix C.

Coding Procedures

The following information was coded: (a) bibliographical information (e.g., authors; year and title of the study; type of publication; country in which the study was performed), (b) study sample (number of participants in each condition and being primary school students, secondary school students or university students), (c) assignment of participants to conditions (randomization on individual level or not) (d) number of intervention sessions, (e) text characteristics (narrative or informative; text only, text and pictures or text and video or animations; text order being linear or not), (f) feedback characteristics (timing of the feedback, type of feedback, means of providing feedback), (g) type of outcome measures (direct post-test or delayed post-test; multiple choice questions, open questions, recall/retelling tasks, ordering pictures or text elements), and (h) information processing level that was aimed at in the outcome measures (factual knowledge versus inferential knowledge). Timing of feedback was coded as during reading or directly after reading. The type of feedback was coded in the following three categories: knowledge-of-response (KOR), knowledge-of-correct-response (KCR) or elaborated feedback (EF). The means of providing feedback was coded in two categories: computer-delivered feedback and non-computer-delivered feedback.

All studies were coded by the first author and undergraduate students in Education and Child studies, who worked in pairs and were trained in coding beforehand by the first author. Two students coded the same report and had to reach consensus on each coding category. Inter-coder reliability was calculated between the scores that the two undergraduate students agreed on and those of the first author. Inter-coder reliability was on average $\kappa = .77$ ($SD = .13$), ranging from $\kappa = .58$ for timing of feedback to $\kappa = 1.00$ for the number of intervention sessions and delayed post-test measures. In case of disagreements, the first author searched through the intervention reports an extra time and made a final coding decision. If essential information on the intervention was not reported, we looked up the software on the Internet (in cases where the used software was available online), and/or we tried to contact the authors by email (in case of reports that were published less than 10 years ago, i.e., published after 2009). Also, if reports did not include sufficient statistical information, we tried to contact authors for studies that were published after 2007. We excluded older studies that did not include sufficient statistical information and/or treated missing information on the interventions as missing data.

If results were reported for subgroups of participants, we separately coded each contrast with the corresponding control group and calculated effect sizes for the different subgroups. If a study included several feedback conditions, we also separately coded each contrast with the corresponding control condition and calculated effect sizes for the different feedback conditions. If a study included one control group and more than one eligible feedback condition, we divided the number of participants in the control group by the number of feedback conditions to prevent us from including the same participants from the control group more than once in the analyses (for a similar approach see Takacs et al., 2014, 2015; Mol et al., 2008, 2009). One feedback condition was chosen in studies where splitting the control group resulted in a group of less than 10 participants. In these cases, the most elaborate feedback condition was included in the meta-analysis (e.g., the KCR condition in Adams & Strickland, 2012).

Statistical Methods

The dependent variable in the present meta-analysis was the difference in mean scores on learning from text between the feedback condition and the control condition. In order to standardize for the use of different scales and types of outcome measures and to correct for possible biases due to the small sample sizes in most of the primary studies, Hedges' g was calculated for the difference between these two conditions. A positive effect size indicated that the participants in the feedback condition performed better on learning from text than participants in the control condition. If available, raw means and standard deviations of the post-test scores and, if applicable, delayed post-test scores were used to calculate Hedges' g. However, if these weren't available, gain scores, frequency distributions, t-test statistics or F-statistics were used to calculate Hedges' g. Comprehensive Meta-Analysis software, Version 2.0 (Borenstein et al., 2005) was used to calculate the effect sizes and for further meta-analytical procedures. Effect sizes for all measurement instruments and tasks were inspected for outliers (i.e., effect sizes with a standardized residual exceeding ± 3.29; Tabachnick & Fidell, 2007). In case of the presence of outliers, these were winsorized into a value .001 higher or lower than the highest or lowest non-outlying effect size. In case of more than one outlier per outcome measure, the rank order of effect sizes was maintained by adding .001 to each next outlier. Subsequently, to account for dependency among multiple effect sizes within a study, if two or more outcome measures for learning were reported in a study, the effect sizes for these different outcome measures were averaged.

In order to take into account the heterogeneity in interventions, study designs and study samples among the studies included in the present meta-analysis, the random effects model was used for combining the effect sizes of different studies and to calculate the 95% confidence intervals of the combined effects (Borenstein et al., 2009; Lipsey & Wilson, 2001; Raudenbush, 2009; Viechtbauer, 2007). When combining the effect sizes for different studies, effect sizes were weighted by their inverse variance, so that studies with larger samples and smaller standard errors had a greater weight on the mean effect size (Lipsey & Wilson, 2001; Shadish & Haddock, 2009). Heterogeneity of effect size was estimated based on the Q-statistic, with a significant Q indicating that the variability among effect sizes was larger than may be expected based on sampling error only (Lipsey & Wilson, 2001).

To investigate the robustness of the effect of feedback on learning from text, we used three indicators for the presence of publication bias in our data. First, we checked our results for publication bias graphically by inspecting the funnel plot with the effects of feedback on learning from text for all studies. Asymmetry in this plot could indicate publication bias, the overrepresentation of significant and large effects in the literature, because those are more likely to get published (Borenstein et al., 2009). In case of publication bias, Duval and Tweedie's (2000) Trim and Fill procedure was used to adjust for this publication bias. Second, we checked the classic fail-safe *N* to investigate how many null-effects would be needed to turn a significant effect into a non-significant one. In line with Rosenthal's criterion (Rosenthal, 1979), we considered an effect robust if the fail-safe *N* exceeded $5k + 10$ (with *k* representing the number of study contrasts). Third, we checked how many effect sizes would be required according to Orwin's fail-safe *N* to reduce the combined effect to a value lower than 0.01 (Orwin, 1983; for a similar procedure see Davis, 2018).

To answer our research questions, moderator analyses were performed using the random effects model to test for differences between the effects of timing of feedback, different types of feedback and the means by which feedback was provided to the reader. Regarding the small number of studies in some categories ($k < 20$) and the fact that studies are not equally distributed among categories, a random effect model was used that assumed a common among-studies variance component (Rubio-Aparicio et al., 2017). Only variables that included at least four contrasts in each cell were used for moderator analyses. A moderator was considered significant if the $Q_{between}(df)$ statistic was significant.

Results

Descriptive Analyses

The literature search for the present meta-analysis resulted in 60 studies, published between 1962 and 2016 (see Appendix D for references to all publications). The studies contained 104 contrasts of reading with feedback conditions versus control conditions of reading without receiving feedback. One contrast was described in a technical report, eight contrasts in conference proceedings, 21 contrasts in dissertations and 74 contrasts in journal articles. The majority of the contrasts, 85, were conducted in the USA, six contrasts were conducted in Germany, five in Spain, four in Taiwan, two in Canada, one in Iran and one in the UK. In twelve of the contrasts participants were not assigned randomly to conditions or random assignment was performed on classroom/group level, in the remaining 92 contrasts participants were randomly assigned to conditions. The contrasts included a total of 178 effects of feedback on outcome measures for learning from text, ranging from -0.91 to 2.41. All contrasts and effects are shown in Appendix E. In total 6,124 participants, ranging from primary school children to university students, were included in the contrasts. The average sample size per contrast was 59.46 (*SD* = 48,64). Information about the level of information processing (factual or inferential knowledge, i.e., superficial or deep level information processing) at which the questions in the immediate and delayed post-tests were targeted was hardly ever reported, which made it impossible to code this information and include it in the present meta-analysis.

Inspection of all effect sizes showed three outlying positive effects for learning from text. These effect sizes were winsorized into a value of respectively 0.001, 0.002 and 0.003 higher than the highest non-outlying effect size in order to maintain the rank order of effect sizes.

The Effect of Feedback on Learning From Text

To investigate the effect of feedback on learning from text, we combined the effect sizes for all contrasts, resulting in an average effect of $g+ = 0.35$ ($k = 104$, $SE = 0.05$, 95% CI = [0.25, 0.46], $p < .001$). Feedback had a positive effect on learning from text. To test the reliability of the overall effect of feedback, we inspected for publication bias and the robustness of the effect. The funnel plot used to inspect for publication bias showed a symmetric pattern of effect sizes and Duval and Tweedie's trim and fill procedure indicated that no extra studies had to be imputed to obtain a symmetric distribution of effects. Additionally, the classic fail-safe *N* indicated that 3,662 contrasts with a null-effect would be needed to turn the significant effect of feedback on learning from text into a non-significant one. According to Orwin's fail-safe *N*, 2,818 contrasts with a null-effect would be needed to reduce the overall effect size to a value lower than 0.01. Based on these statistics, we concluded that the effect of feedback on learning from text was reliable and robust.

To test for other possible biases, moderator analyses were performed for country, assignment of participants to conditions (random on an individual level vs. not random or random on classroom/group level), subject design (within vs. between), text type (narrative, informative or both), instructional design (text only, text + pictures or text + video/animations), type of outcome measure (open question, MC-questions or both) and subject's age group (elementary school, secondary school, students) and meta-regression analyses were performed for publication year, frequency of feedback relative to text length and the number of intervention sessions. Except for subject design and age group, none of the regression models or moderators were significant. On average, studies with a within-subject design showed a larger effect of feedback on learning from text ($g+ = 0.93$, $k = 4$, $SE = 0.26$, 95% CI = [0.41, 1.45], $p < .001$) than studies with a between-subject design ($g+ = 0.33$, $k = 100$, $SE = 0.05$, 95% CI = [0.23, 0.43], $p < .001$), $Q_{between}(1) = 4.95$, $p = .03$. However, only four contrasts came from within subject studies, so we did not take this difference into account in further analyses. For age group, we found a significant difference in effects in studies from different age populations, $Q_{between}(2) = 14.54$, $p = .001$. Neither for elementary school students ($g+ = 0.32$, $k = 5$, $SE = 0.23$, 95% CI = [-.12, 0.76], $p = .16$) nor for secondary school children ($g+ = 0.09$, $k = 34$, $SE = 0.09$, 95% CI = [-0.09, 0.26], $p = .32$) we found a significant effect of feedback on learning from text. For university students however, there was a significant positive effect of feedback ($g+ = 0.50$, $k = 65$, $SE = 0.07$, 95% CI = [0.37, 0.63], $p < .001$). A significantly larger proportion of the studies with secondary school students (79.1%) used feedback during reading in the interventions than was the case for the studies with university students (61.5%, $\chi^2 = 3.80$, $p = .05$) which might explain the lack of effect but other explanations might hold as well which will be elaborated on in the discussion. In order to increase power to conduct the proposed moderator analysis, the low number of studies with primary school students, and to maintain the focus of the present meta-analysis on variables regarding the design of the feedback, results will not be separately reported for the three age groups.

For post-tests of learning from text that were applied immediately after the intervention was completed, the average effect of feedback was $g+ = 0.32$ ($k = 102$, $SE = 0.05$, 95% CI = [0.22, 0.41], $p < .001$), for delayed post-test measures the average effect was $g+ = 0.30$ ($k = 23$, $SE = 0.14$, 95% CI = [0.04, 0.57], $p = .03$). Time between the interventions and delayed post-test measures varied from one to ten weeks. Because for most of the study contrasts (77.9%) only immediate post-test scores were reported, we conducted further analyses only on the immediate post-test scores of learning from text. As the effect of feedback on immediate post-tests was heterogeneous ($Q(101) = 324.78$, $p < .001$), we conducted moderator analyses to test for differences in the effects of feedback concerning feedback timing, types of feedback and the means by which feedback was provided.

Timing of Feedback

For two contrasts, the reports did not provide enough information to determine if feedback was provided to participants during reading or directly after reading the text. Out of the remaining hundred contrasts, in 69 studies the effect of feedback during reading and 31 the effect of feedback directly after reading on learning from text was tested. Moderator analysis showed that feedback directly after reading ($g+ = 0.46$, $k = 31$, $SE = 0.09$, 95% CI = [0.28, 0.64], $p < .001$) was more effective in enhancing learning from text than feedback during reading ($g+ = 0.23$, $k = 69$, $SE = 0.06$, 95% CI = [0.11, 0.34], $p < .001$), $Q_{between}(1) = 4.32$, $p = .04$ (see Table 1). Both the effects of feedback during reading and feedback directly after reading on learning from text were heterogeneous (see Table 1). Therefore, moderator analyses for the effects of different types of feedback were performed for contrasts studying feedback during reading and directly after reading separately.

Different Types of Feedback

Feedback Directly After Reading

A moderator analysis showed that feedback type was a significant moderator of the effect of feedback on learning from text ($Q_{between}$ (2) = 7.38, p = .03), see Table 2 and Figure 1. If provided directly after reading, elaborate feedback (EF), showed the largest positive effect on learning from text ($g+$ = 0.92, k = 6, SE = 0.25, 95% CI = [0.42, 1.42], p < .001). Knowledge-of-correct-response feedback (KCR) had a moderate positive effect on learning from text ($g+$ = 0.61, k = 13, SE = 0.18, 95% CI = [0.27, 0.96], p < .001). Knowledge-of-response feedback (KOR) had no significant effect on learning from text ($g+$ = 0.14, k = 12, SE = 0.18, 95% CI = [-0.21, 0.48], p = .45). All effects were heterogeneous. The difference in effects between EF and KCR did not reach significance ($Q_{between}$ (1) = 0.74, p = .39), but EF was more effective than KOR in enhancing learning from text ($Q_{between}$ (1) = 8.63, p = .003). Also, KCR was more effective than KOR ($Q_{between}$ (1) = 4.21, p = .04).

Feedback During Reading

A moderator analysis on feedback type for contrasts testing the effects of feedback during reading on learning from text showed that the effects of different types of feedback did not differ significantly from each other ($Q_{between}$ (3) = 2.86, p = .41), see Table 2 and Figure 1. Elaborate feedback had a positive, small, and significant effect on learning from text ($g+$ = 0.25, k = 24, SE = 0.25, 95% CI = [0.08, 0.42], p < .01), as did feedback containing KOR ($g+$ = 0.39, k = 11, SE = 0.39, 95% CI = [0.14, 0.64], p < .01). Neither feedback containing KCR (g+ = 0.14, k = 30, SE = 0.14, 95% CI = [-0.02, 0.30], p = .08) nor feedback that alternately contained KOR and KCR in the same study (g+ = 0.22, k = 4, SE = 0.22, 95% CI = [-0.20, 0.64], p = .31) had a significant effect on learning from text. All effects were heterogeneous.

Means of Providing Feedback

Feedback Directly After Reading

Non-computer-delivered feedback after reading did not have a significant effect on learning from text ($g+$ = 0.19, k = 8, SE = 0.24, 95% CI = -0.27, 0.65], p = .42). Computer-delivered feedback had a medium positive effect on learning from text ($g+$ = 0.59, k = 23, SE = 0.14, 95% CI = 0.32, 0.85], p < .001). This effect was heterogeneous (see Table 3). The difference between the two means of providing feedback was not significant ($Q_{between}$ (1) = 2.15, p = .14).

Feedback During Reading

Non-computer-delivered feedback during reading did not have a significant effect on learning from text ($g+ = 0.10$, $k = 28$, $SE = 0.08$, 95% CI = [-0.07, 0.26], $p = .24$), but computer-delivered feedback had a small significant effect on learning from text ($g+ = 0.30$, $k = 41$, $SE = 0.07$, 95% CI = [0.18, 0.43], $p < .001$). Both effects were heterogeneous (see Table 3). The difference between the two means of providing feedback was marginally significant ($Q_{between} = 3.78$, $p = .05$).

Discussion

The aim of the present study was to investigate the effects of feedback on learning from text in conventional readers. Based on the 104 contrasts included in the meta-analysis it can be concluded that feedback positively influences learning from text. However, the magnitude of the effect of feedback on learning from text ($g+ = 0.35$; $d = 0.36$) is smaller than the effects of feedback found in most of the previous meta-analyses on the effects of feedback on learning in general (e.g., Kluger & DeNisi, 1996; $d = 0.41$; Azevedo & Bernard, 1995; $d = 0.80$; Hattie, 2012; $d = 0.79$). Van der Kleij et al. (2015) concluded in their meta-analysis on the effects of feedback in computer-based learning environments that feedback was particularly effective for promoting mathematics performance or problem-solving skills, noting that feedback in these subject areas was mostly elaborate feedback, compared to instruction in social sciences and language, subject areas in which feedback was more often KOR- of KCR-feedback. The variation in the types of feedback included in the present meta-analysis could explain why the average effect we found in the present meta-analysis is lower. Another explanation for the lower average effect of feedback on learning from text compared to the effects found in other meta-analyses may be based on the mode of presentation of the feedback. According to the dual coding theory (Clark & Paivio, 1991), information can be processed in two different modes; verbal and non-verbal. Information in one mode can only be processed sequentially, information in two modes can be processed in parallel or simultaneously and these two modes can form interconnections which may strengthen information processing. Although Shute (2008) advised, based on her meta-analysis on the effects of feedback on learning, to use multiple modes (i.e., an audio and a visual mode; see also Clark & Paivio, 1991) for the presentation of feedback to prevent creating a cognitive overload due to modality effects (i.e., presenting both the learning material and feedback in the same modality; see also Moreno & Mayer, 1999), in the present meta-analyses in almost all studies (96.2%) feedback was presented only as text. In other words, presenting both the text and feedback in the same mode (i.e., both as written text) increases the cognitive load of the reading task, which may explain the difference between the effects of feedback on learning from text found in the present meta-analysis and the effects of feedback on more general outcomes found in previous meta-analyses.

The effect of feedback on learning from text found in the present meta-analysis was heterogeneous. The broad range of effect sizes found in the primary studies included in the present meta-analysis is in line with Hattie's (2012) conclusion that feedback on average has a positive influence on learning, but that much variation exists between studies. In order to explain the variation in effects among studies, we looked at three characteristics of the design of the feedback provided in the primary studies: (a) timing of feedback, (b) different types of feedback and (c) the means of providing feedback.

Timing of Feedback

The present meta-analysis shows that feedback is most effective to support learning from text if provided directly after reading, instead of during reading. This finding is in line with Sweller's (1994) cognitive load perspective and the results of a study by Subrahmanyam et al. (2013) showing that multitasking results in less efficient processing of a text. In line with this idea, Shute (2008) previously formulated a guideline for designing feedback to enhance learning, stating that feedback should not interrupt the learner when he or she is actively engaged in a task, because this could impede learning. As we might infer from the results of the present-meta-analysis, feedback during reading seems to interrupt the natural reading process and forces the reader to multitask between the processing of the text and the questions or tasks and subsequent feedback during the text thus placing extra load on working memory, also called the split-attention effect (Sweller, 1994). This conclusion is comparable to that of a recent meta-analysis on the effect of technology-enhanced storybooks for young children, who were not conventional readers, showing that interactive options in an e-book do not promote comprehension of the story, but seem to distract children from the storyline (Takacs et al., 2015). From the present meta-analysis it appears that even for conventional readers interaction during reading seems to impede learning from a text, even if this interaction is relevant to understand the information in the text, which is not the case in most interactions in technology-enhanced storybooks for young children. According to Sweller et al. (1998), one could argue that providing feedback results in an increased interactivity between elements (i.e., information from the text, the questions or tasks and the feedback) that has to be processed, which results in an increased cognitive load. As a consequence, less working memory capacity is available to integrate the information into a coherent mental model of the text, which impedes learning from the text.

Brangert-Drowns et al. (1991) proposed in their five-stage model of the effects of feedback on learning that feedback should trigger the reader to evaluate and adjust task-relevant knowledge. It seems that, in contrast to the guidelines of Mullet and Marsh (2016) stating that feedback should function as a 'forget-cue', the best moment for reflecting on task-relevant knowledge (i.e., a reader's mental model of the text), is not necessarily directly after a possible error occurred. The present meta-analysis showed that it is more effective to first build a complete mental model of a text and subsequently evaluate and adjust the mental model based on questions or tasks followed by feedback after the text. The increased cognitive load caused by questions and feedback during reading appears to interfere with the construction of a mental model of the text.

The finding that feedback during reading less effectively facilitates the evaluation and adjustment of the reader's mental model of a text than feedback after reading also fits the *imperfect mental model view* of Chi (2000). Chi states that, as a consequence of differences in readers' pre-existing mental models that are based on individuals' background knowledge and experience, the way and pace at which readers build mental models of sentences or parts of a text and detect possible flaws in the mental model are not the same for everyone. In line with this reasoning, the moment at which a question or task and subsequent feedback regarding a specific part of the text could be most helpful differs per reader. Providing questions or tasks and feedback at the end of a text enables readers to first engage in reading and mental model building in their own pace and manner. Subsequently, feedback can help the reader to revise the mental model where needed based on the information provided in the feedback messages.

Different Types of Feedback

The amount of information that feedback messages contain also appears to influence the effect that feedback has on learning from text. As a consequence of the aforementioned differences in the effects of timing of feedback, we investigated the effects of different types of feedback separately for feedback provided during and feedback provided directly after reading. If provided after reading, EF and KCR appeared to be most effective, more so than KOR, in promoting learning from text. Although the largest effect was found for EF ($g+ = 0.92$), this was not significantly different from the effect of KCR on learning from text ($g+ = 0.61$). This is possibly due to the low number of studies containing elaborate feedback ($k = 6$), which reduces the statistical power of the comparison. The large positive effect of EF on learning from text is in line with the results of previous meta-analyses showing that elaborated feedback is particularly effective for promoting higher order learning tasks and deep learning (Hattie & Timperley, 2007; Jaehnig & Miller, 2007; van der Kleij et al., 2015). As described in the model of Kulhavy and Stock (1989) an elaborate feedback message not only contains verification, but also the correct answer and some instructional information (e.g., explaining why an answer is correct or repeating the part of a text that contained information regarding the correct answer) or extra-instructional information (e.g., providing information, examples or analogies that were not present in the original learning material). This instructional- or extra-instructional information can help the reader to better process and understand the learning material by helping the reader to improve the interconnections between the building blocks of the mental model instead of only knowing that the mental model is correct or incorrect, which is in fact the only message in KOR or (to a lesser degree) in KCR (see also Hattie & Timperley, 2007). In line herewith, Kluger and DeNisi (1996) stated in their Feedback Intervention Theory (FIT) that the extra information in the feedback message provides the learner with guidance on how to invest additional effort to succeed in task performance: in the case of learning from text to reach or restore completeness and coherence of the mental representation of the text.

For studies in which feedback was provided during reading, no differences were found between the different types of feedback. So, although it appears to be most effective to provide elaborated feedback after reading, during reading only stating if an answer was correct or not appeared to be as effective as elaborate feedback. This contrast could possibly be explained by the idea that feedback during reading hinders the natural reading process, and, therefore, the building of a mental model of the text. Although elaborate feedback contains more information than KOR, it takes more time and effort to process an elaborate feedback message than a simple *right* or *wrong*. In other words, the extraneous cognitive load (Sweller, 1994) of an elaborate feedback message is larger than of KOR. As a consequence, elaborate feedback might interrupt reading more than KOR, which impedes learning from the extra information in the feedback message and from the text. Another explanation for the equal effectiveness may come from a phenomenon called the redundancy effect (Sweller et al., 1998). If the feedback message presents information that is closely similar to the information in the text, in particular for better skilled readers, this information may be redundant, creating an unnecessary load on the limited working memory capacity of the reader. At the same time, for less skilled readers this information could be a valuable help to integrate the information from the text into a coherent mental model. However, in the present meta-analysis we were not able to take into account differences in students' reading levels as these were often not reported.

The Means of Providing Feedback

The present meta-analysis shows that computer-delivered feedback has a positive effect on learning from text, but that non-computer-delivered feedback has not. Computer-delivered feedback automatically appears when a reader has completed a question or task whereas it takes more time and effort to process a non-computer-delivered feedback message by, for example, retrieving papers from a separate folder, scratching answer forms, using latent image pens or removing flaps that cover the feedback. The difference between these two means of providing feedback seemed especially prominent for feedback presented during reading. These findings are in line with our hypothesis that the time and effort needed to uncover the non-computer-delivered feedback before it can be processed, places a higher demand on working memory capacity (split-attention effect; Sweller, 1994). This demand on working memory seems especially detrimental for learning when feedback is provided during reading, because this form of feedback appeared to already place a burden on working memory.

Limitations and Future Research

In the present meta-analysis we focused on the effects of feedback on learning from text in conventional readers. Nevertheless, this is still a heterogeneous group containing elementary school children, secondary school children and students, all with different levels of reading skills and meta-cognitive skills and reading texts at different levels of complexity. Unfortunately, the fact that a large majority of studies were conducted with college students and university students, the low number of studies with elementary school children and the large overlap between age groups of participants and the timing of feedback, made it impossible to draw strong conclusions about the effects of feedback for different age groups. Nevertheless, the difference in effects of feedback on learning from text between secondary school students and university school students suggest a developmental aspect in the processing of and profiting from feedback. Secondary school children profited less from feedback than university students. Based on the model that the relation between reading achievement and meta-cognitive skills are bi-directional (Edossa et al., 2019) it is plausible that younger readers (i.e., secondary school students) have less well developed meta-cognitive skills, thus responding differently on feedback compared to more proficient readers (i.e., university students). Alternatively, the difference in the effects of feedback on learning from text between secondary school students and university students may arise from differences in students conceptions of feedback. Peterson and Irving (2008) found in their study on students conceptions of assessment and feedback that most secondary school students report that they want feedback that helps them improving learning, but at the same time that many also admit to ignore or forget the given feedback, because they are mainly focused on the summative assessment of their performance (i.e., their grade). As a consequence, they don’t invest to use the feedback to actively improve their performance. From a motivational perspective, differences in the effects of feedback on learning from text between secondary school students and university students may also be explained by the intrinsic motivation of students to perform well in the reading tasks during the feedback interventions. Wigfield et al. (2016) argued that a decrease in reading motivation in secondary school children could partly be explained by a lack of students' believe that what they are learning is relevant. In studies with university students, one could expect students to be much more aware of the reasons why the learning material they are reading is relevant than for secondary school children. This could have consequences for both students' conceptions of feedback and the performance of the students on the reading tests.

Additionally, the groups of elementary school children and secondary school children were probably much more heterogeneous regarding their reading levels than university students. Unfortunately, we were not able to investigate possible differences in the effects of feedback for high or low performing readers. Because low performing or beginning readers need to spend more time and effort to decode words and sentences, less working memory capacity is left to build a mental model of a text (Cain, 2010). As a consequence, these types of readers may also have less working memory capacity left to process questions and the subsequent feedback. This may have implications for the way feedback should be designed for these groups of readers. For example, low performing readers may need other or extra information in feedback messages than high performing readers (i.e., redundancy effect; Sweller et al., 1998). Future research investigating the effects of feedback on learning from text separately for readers of different ages and low and high performing readers could help to gain insight in this matter.

In the present meta-analysis we categorised feedback as KOR, KCR or elaborated feedback (EF). However, as described in the models of Kulhavy and Stock (1989) and Hattie and Timperly (2007), the information in these elaborated feedback messages could greatly vary. To develop a more thorough understanding of what kind of elaborated feedback messages may be most effective in enhancing learning from texts, future research could focus on differences in the effects of different types of elaborated feedback messages (i.e., containing only instructional information (i.e., a repetition of information that is already given in the text) or both instructional and extra-instructional information (i.e., new information or explanations in addition to those already provided in the text); Kulhavy & Stock, 1989; or process-level feedback compared to self-regulation-level feedback; Hattie & Timperley, 2007). According to Kluger and DeNisi (1996) different types of feedback messages may elicit different strategies or learning processes by the learner to reduce the gap between current performance (i.e., in the case of learning from text, the current understanding or mental model of the text) and the goal of a task (i.e., in the case of learning from text, a complete and coherent mental model of the text). The deployment of these strategies may influence the effectiveness of the feedback on learning from text.

Alternatively, in future research the effects of different types of feedback on different levels of information processing could be tested. Information about the level of information processing at which the questions in the direct and delayed post-tests were targeted was hardly reported, which made it impossible to test differences in the effects of feedback on different levels of learning. Learning from or the understanding of a text can be measured on different levels, with an increase in difficulty ranging from questions about facts or sentence-based inferences (i.e., a text-based model or propositional model of the text) to relations between information in sentences among the text and between information in the text and background knowledge (i.e., a situation-model of the text; see Kintsch, 1988). In line with the idea that elaborate feedback would be especially effective to support higher-order learning or deep processing of information, contrary to KOR or KCR, types of feedback that would be more suitable to support knowledge on a surface level (Hattie & Timperley; 2007; van der Kleij et al., 2012), it could be tested in future studies if KOR and KCR would be more effective in supporting learning from text on the level of a superficial, text-based model and if EF would be more effective to promote deep processing of the information in the text.

In the present meta-analysis 74 out of 104 contrasts included interventions of a single session and another 11 contrasts included intervention of at maximum five sessions. In other words, a large majority of the studies was focused on how to provide support 'on the job', i.e., while reading text, to enhance learning from text compared to testing the transferability of strategies targeted by the feedback. In line with this, the majority of the post-test measures used in the primary studies tested the understanding of the content of the texts that were read during the interventions. Hattie and Timperley (2007) argued that feedback on a processing level could help learners to develop effective error-detection skills, information search skills and the use of task strategies, which they can use in new situations, where feedback is not provided by an external agent (e.g., by a computer, teacher or peer), to enhance learning. These metacognitive processes are essential for developing comprehension monitoring skills, that should be used during reading to make meaning from a text (Baker & Brown, 1984).

Finally, in the present meta-analysis we discussed the results from a cognitive load perspective. However, the question of how to determine cognitive load is difficult because of its multidimensional character and the complex interrelationships between performance, cognitive load, and mental effort (see Sweller et al., 1998). In the current meta-analysis it was not possible to test the interaction between different sources of cognitive load (for example the complexity of the information in both the text and feedback messages, modality of presenting feedback, the use of effortful learning processes compared to more automatic learning processes), due to the fact that information with regard to these resources was not reported in the original papers.

Conclusion

The present meta-analysis, including 104 contrasts of reading texts with and without receiving feedback, adds to the field by its narrower scope by specifically testing effects of feedback on learning from texts compared to learning in general in previous meta-analyses (e.g., Hattie, 2012; Shute, 2008; van der Kleij et al., 2015). Also, we only included studies in which there was a direct comparison between a reading with versus a reading without feedback condition. The results of our meta-analysis show that providing feedback during or directly after reading has a positive effect on learning from text. Promoting learning from text appears to be most effective when feedback is presented after reading a text and at least contains the correct answer or is elaborate. Feedback during reading is less effective, probably because it hinders the natural reading process, thereby placing too heavy demands on working memory (i.e., split-attention effect). Additionally, computer-delivered feedback is more beneficial for learning from text than non-computer-delivered feedback. When developing or choosing (educational technologies for) instructional strategies to support learning from text, one should keep in mind that it seems best to minimally interrupt the reading process (i.e., placing minimal load on the limited working memory capacity of the reader), and to help the reader evaluate and, if necessary, revise the mental model of a text with the help of questions or tasks and subsequent elaborate computer-delivered feedback directly after reading.

Table 1

Effects of Feedback During Reading and Directly After Reading on Learning From Text

Moderator variable	Subgroup of studies	*k*	Average effect size (g+)	*SE*	95% CI	Heterogeneity statistics
Timing of feedback	After reading	31	0.46*	0.09	[0.28, 0.64]	130.58, $p < .001$
	During reading	69	0.23*	0.06	[0.11, 0.34]	167.73, $p < .001$

Note: $Q_{between}$ (1) = 4.32, $p = .04$; * $p <.001$

4

Table 2

Effects of Different Types of Feedback on Learning From Text, Separated for Timing of Feedback

Timing of feedback	Type of feedback	*k*	Average effect size (*g*+)	*SE*	95% CI	*p*	Heterogeneity statistics (Q_{within})
After reading	KOR	12	0.14	0.18	[-0.21, 0.48]	.45	6.65, *p* = .65
	KCR	13	0.61	0.18	[0.27, 0.96]	<.001	65.92, *p* < .001
	Elaborate	6	0.92	0.25	[0.42, 1.42]	<.001	33.61, *p* <.001
During reading	KOR	11	0.39	0.13	[0.14, 0.64]	< .01	29.67, *p* = .001
	KOR + KCR	4	0.22	0.21	[-0.20, 0.64]	.31	1.12, *p* = .77
	KCR	30	0.14	0.08	[-0.02, 0.30]	.08	66.74, *p* < .001
	Elaborate	24	0.25	0.09	[0.08, 0.42]	< .01	59.32, *p* < .001

Note: KOR = knowledge-of-response feedback, KCR = knowledge-of-correct-response feedback, EF = elaborate feedback; Moderator statistics for feedback type after reading: $Q(2) = 7.38$, $p = .03$; Moderator statistics for feedback type during reading: $Q(3) = 2.86$, $p = .41$

Table 3

Effects of the Means of Providing Feedback on Learning From Text, Separated for the Timing of Feedback

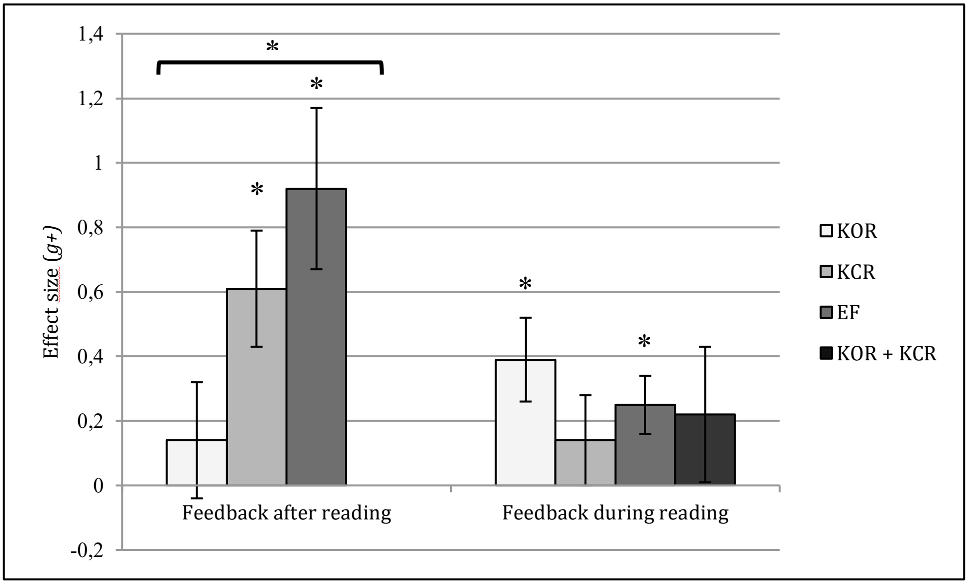
Timing of feedback	Means of providing feedback	*k*	Average effect size ($g+$)	*SE*	95% CI	*p*	Heterogeneity statistics (Q_{within})
After reading	Non-computer-delivered	8	0.19	0.24	[-0.27, 0.65]	.42	5.71, p = .57
	Computer-delivered	23	0.59	0.14	[0.32, 0.85]	< .001	121.71, p < .001
During reading	Non-computer-delivered	28	0.10	0.08	[-0.07, 0.26]	.24	61.98, p < .001
	Computer-delivered	41	0.30	0.07	[0.18, 0.43]	<.001	98.79, p < .001

Note: Moderator statistics for the means of providing feedback after reading: $Q(1) = 2.15$, $p = .14$; Moderator statistics for the means of providing feedback during reading: $Q(1) = 3.78$, $p = .05$

4

Figure 1

Effects of Different Types of Feedback on Learning From Text, Separated for Feedback Timing.



Note: KOR = knowledge-of-response feedback; KCR = knowledge-of-correct-response feedback; EF = elaborate feedback. If feedback was provided after reading, EF was more effective than KOR or KCR to support learning from text. If feedback was provided during reading, no differences were found between the effects of different types of feedback on learning from text. Error bars represent the 95% confidence intervals of the effect sizes. **p* , .05. ** *p* < .01. ****p* < .001.

Appendix A

Combinations of Terms Used for Literacy Search

- book* AND feedback OR book* AND agent OR book* OR book* AND cai OR book* AND tutor OR book* AND feedback AND agent OR book* AND feedback AND avatar OR book* AND feedback AND cai OR book* AND feedback AND model* OR book* AND feedback AND scaffold* OR book* AND feedback AND tutor* OR book* AND interactive AND agent OR book* AND interactive AND avatar OR book* AND interactive AND cai OR book* AND interactive AND model* OR book* AND interactive AND scaffold* OR book* AND interactive AND tutor*
- e-book AND agent OR e-book AND avatar OR e-book AND cai OR e-book AND feedback OR e-book AND interactive OR e-book AND model* OR e-book AND scaffold* OR e-book AND tutor*
- electronic book AND agent OR electronic book AND avatar OR electronic book AND cai OR electronic book AND feedback OR electronic book AND interactive OR electronic book AND model* OR electronic book AND scaffold* OR electronic book AND tutor*
- literacy AND agent OR literacy AND avatar OR literacy AND cai OR literacy AND feedback OR literacy AND interactive OR literacy AND model* OR literacy AND scaffold* OR literacy AND tutor*
- literacy AND books AND agent OR literacy AND books AND avatar OR literacy AND books AND cai OR literacy AND books AND feedback OR literacy AND books AND interactive OR literacy AND books AND model* OR literacy AND books AND scaffold* OR literacy AND books AND tutor*
- literacy AND feedback AND agent OR literacy AND feedback AND avatar OR literacy AND feedback AND cai OR literacy AND feedback AND model* OR literacy AND feedback AND scaffold* OR literacy AND feedback AND tutor*
- literacy AND interactive AND agent OR literacy AND interactive AND avatar OR literacy AND interactive AND cai OR literacy AND interactive AND model* OR literacy AND interactive AND scaffold* OR literacy AND interactive AND tutor*
- literacy AND storybooks AND agent OR literacy AND storybooks AND avatar OR literacy AND storybooks AND cai OR literacy AND storybooks AND model* OR literacy AND storybooks AND scaffold* OR literacy AND storybooks AND tutor*
- reading AND books AND agent OR literacy AND reading AND avatar OR reading AND books AND cai OR reading AND books AND feedback OR reading AND books AND interactive OR reading AND books AND model* OR reading AND books AND scaffold* OR reading AND books AND tutor*

- reading AND feedback AND agent OR reading AND feedback AND avatar OR reading AND feedback AND cai OR reading AND feedback AND model* OR reading AND feedback AND scaffold* OR reading AND feedback AND tutor*
- reading AND interactive AND agent OR reading AND interactive AND avatar OR reading AND interactive AND cai OR reading AND interactive AND model* OR reading AND interactive AND scaffold* OR reading AND interactive AND tutor*
- reading AND storybooks AND agent OR reading AND storybooks AND avatar OR literacy AND storybooks AND cai OR reading AND storybooks AND model* OR reading AND storybooks AND scaffold* OR reading AND storybooks AND tutor* OR storybooks AND cai

Appendix B

References of Searched Handbooks

Barr, R., Kamil, M. L., Mosenthal, P., & Pearson, P. D. (Eds.). (1991). *Handbook of reading: Volume II.* Lawrence Erlbaum Associates.

Bus, A. G. & Neuman, S. B. (Eds.). (2008). *Multimedia and literacy development: Improving achievement for young learners.* Routledge.

Dickinson, D. K. & Neuman, S. B. (Eds.). (2006). *Handbook of early literacy research: Volume II.* The Guilford Press.

Flippo, R. F. & Caverly, D. C. (2000). *Handbook of college reading and study strategy research* (2nd edition). Lawrence Erlbaum Associates.

Hall, N. , Larson, J., & Marsh, J. (Eds.). (2003). *Handbook of early childhood literacy.* SAGE Publications.

Israel, S. E. & Duffy, G. G. (Eds.). (2009). *Handbook of research on reading comprehension.* Taylor & Francis.

Kamil, M. L., Mosenthal, P. B., Pearson, P. D. & Barr, R. (Eds.). (2000). *Handbook of reading: Volume III.* Lawrence Erlbaum Associates.

Kinzer, C. & Verhoeven, L. T. (Eds.). (2007). *Interactive literacy education: Facilitating literacy environments through technology.* Lawrence Erlbaum Associates.

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning.* Cambridge University Press.

McKenna, M. C., Labbo, L. D., Kieffer, R. D. & Reinking, D. (Eds.). (2006). *International handbook of literacy and technology: Volume II*. Lawrence Erlbaum Associates.

McNamara, D.S. (Ed.) (2007). *Reading Comprehension Strategies: Theories, Interventions, and Technologies.* Lawrence Erlbaum Associates.

Morrow, L. M., Rueda, R., & Lapp, D. (2009). *Handbook of research in literacy and diversity.* The Guilford Press.

Neuman, S. B. & Dickinson, D. K. (Eds.). (2001). *Handbook of early literacy research: Volume I.* The Guilford Press.

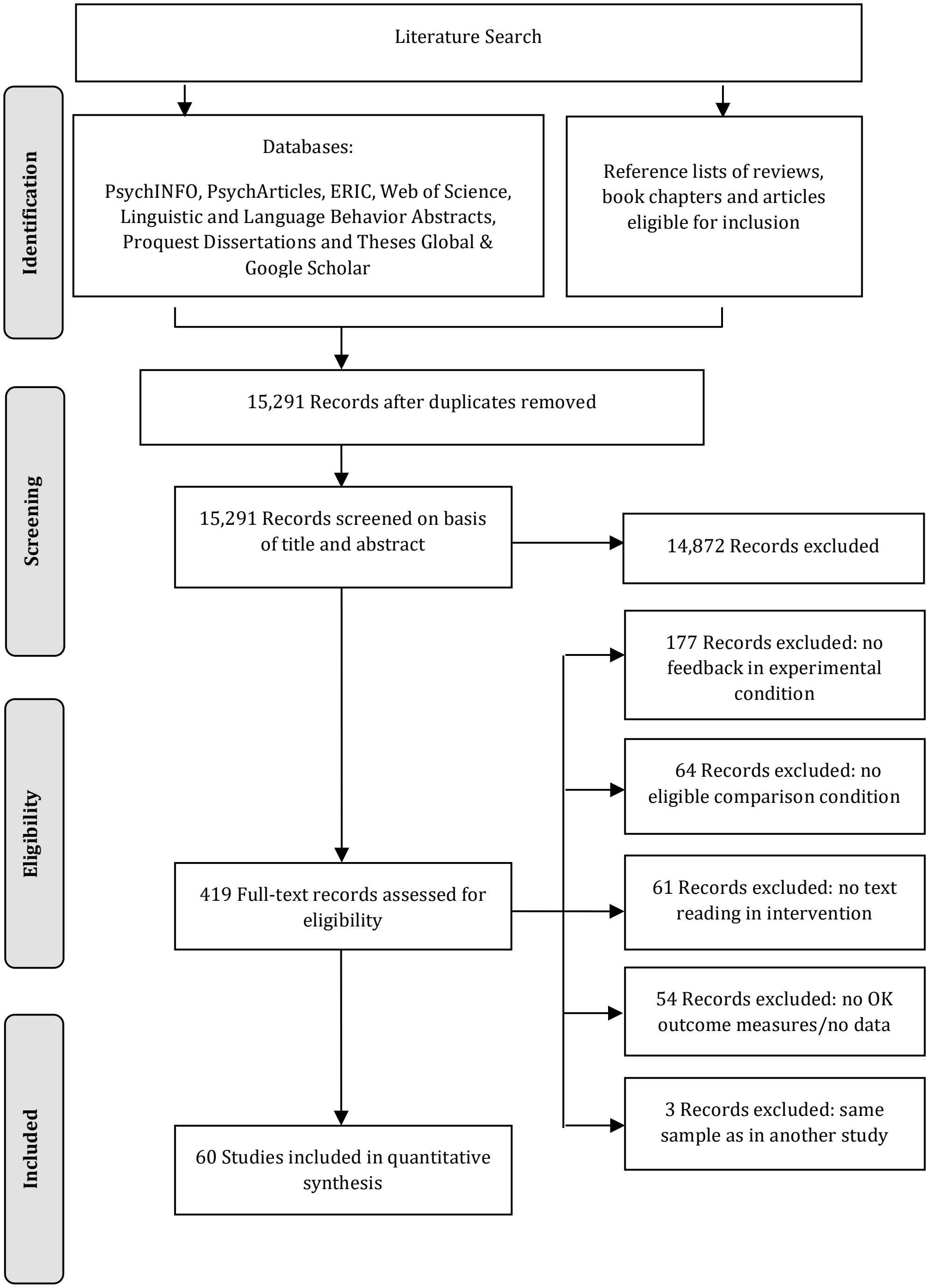
O'Connor, R. E. & Vadasy, P. F. (Eds.). (2011). *Handbook of reading interventions.* The Guilford Press.

Olson, D. R. & Torrance, N. (Eds.). (2009). *The Cambridge handbook of literacy.* Cambridge University Press.

Pearson, P. D., Barr, R., Kamil, M. L., & Mosenthal, P. (Eds.). (2002). *Handbook of reading research: Volume I.* Lawrence Erlbaum Associates.

Appendix C

PRISMA Flow Diagram



Appendix D

References of Studies Included in the Present Meta-Analysis

Adams, R. H., & Strickland, J. (2012). The effects of computer-assisted feedback strategies in technology education: A comparison of learning outcomes. *Journal of Educational Technology Systems, 40*, 211-223. doi:10.2190%2FET.40.2.i

Anderson, R. C., Kulhavy, R. W., & Andre, T. (1971). Feedback procedures in programmed instruction. *Journal of Educational Psychology*, *62*, 148-156. doi:10.1037/h0030766

Anderson, R. C., Kulhavy, R. W., & Andre, T. (1972). Conditions under which feedback facilitates learning from programmed lessons. *Journal of Educational Psychology, 63*, 186-188. doi:10.1037/h0032653

Azevedo, R., Landis, R. S., Feyi-Behnagh, Duffy, M., Trevors, G., Harley, J. M., Bouchet, F., Burlison, J., Taub, M., Pacampara, N., Yeasin, M., Rahman, A. K. M. M., Tanveer, M. I., & Hossain, G. (2012). The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with MetaTutor. In: S. A. Cerri, W. J. Clancey, G. Papadourakis, K. Panourgia (Eds). *Intelligent Tutoring Systems. ITS 2012. Lecture Notes in Computer Science, Vol. 7315* (pp. 212-221). Springer. doi:10.1007/978-3-642-30950-2_27

Butler, A. C., Godbole, N., & Marsh, E. J. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *Journal of Educational Psychology, 105*, 290-298. doi:10.1037/a0031026

Chanond, K. (1988). *The effects of feedback, correctness of response, and response confidence on learner's retention in computer-assisted instruction* (Publication No. 8816422) [Doctoral dissertation, The University of Texas at Austin]. Proquest Dissertations and Theses Global.

Chen, N. S., Teng, D. C. E., & Lee, C. H. (2011). Augmenting paper-based reading activity with direct access to digital materials and scaffolded questioning. *Computers & Education, 57*, 1705-1715. doi:10.1016/j.compedu.2011.03.013

Clariana, R. B., & Koul, R. (2006). The effects of different forms of feedback on fuzzy and verbatim memory of science principles. *British Journal of Educational Psychology, 76*, 259-270. doi:10.1348/000709905X39134

Clariana, R.B., Ross, S.M. & Morrison, G.R. (1992, February 5-9). *The effects of different feedback strategies using computer-administered multiple-choice questions as instruction* [Paper presentation]. The 14th Convention of the Association for Educational Communications and Technology and Sponsored by the Research and Theory Division. Washington D. C., USA.

Clark, K., & Dwyer, F. (1998). Effect of different types of computer-assisted feedback strategies on achievement and response confidence. *International Journal of Instructional Media, 25*(1), 55-63.

Elliot, B. A. (1986). *An investigation of the effects of computer feedback and interspersed questions on the text comprehension of poor readers (computer-assisted instruction)* (Publication No. 303529305) [Doctoral dissertation, Temple University]. ProQuest Dissertations & Theses Global.

Farragher, P. (1980). *Learning environmental science from text aided by a diagnostic and prescriptive instructional strategy* (Publication No. 8024445) [Doctoral dissertation, The Pennsylvania State University). ProQuest Dissertations & Theses Global.

Farragher, P., & Yore, L. D. (1997). The effects of embedded monitoring and regulating devices on the achievement of high school students learning science from text. *School Science and Mathematics, 97*, 87-95. doi:10.1111/j.1949-8594.1997.tb17348.x

Fazio, L. K., Huelser, B. J., Johnson, A., & Marsh, E. J. (2010). Receiving right/wrong feedback: Consequences for learning. *Memory, 18*, 335-350. doi:10.1080/09658211003652491

Feldhusen, J. F., & Birt, A. (1962). A study of nine methods of presentation of programmed learning material. *The Journal of Educational Research, 55*, 461-466. d doi:10.1080/00220671.1962.10882854

Feng, B., & Reigeluth, C. M. (1983). *The effect of three different kinds of feedback: Hint, correct answer, and right/wrong* (IDD&E Working Paper No. 11). Syracuse University, New York School of Education.

Fernald, P. S., & Jordan, E. A. (1991). Programmed instruction versus standard text in introductory psychology. *Teaching of Psychology, 18*, 205-211. doi:10.1207%2Fs15328023top1804_1

Franzke, M., Kintsch, E., Caccamise, D., Johnson, N., & Dooley, S. (2005). Summary Street®: Computer support for comprehension and writing. *Journal of Educational Computing Research, 33*, 53-80. doi:10.2190/DH8F-QJWM-J457-FQVB

Goldbeck, R. A., & Campbell, V. N. (1962). The effects of response mode and response difficulty on programed learning. *Journal of Educational Psychology, 53*, 110-118. doi:10.1037/h0044479

Golke, S., Dörfler, T., & Artelt, C. (2009). Effects of accuracy feedback during a text comprehension test. *Educational & Child Psychology, 26*(3), 30-39. doi:10.1007/s10763-014-9575-4

Golke, S., Dörfler, T., & Artelt, C. (2015). The impact of elaborated feedback on text comprehension within a computer-based assessment. *Learning and Instruction, 39*, 123-136. doi:10.1016/j.learninstruc.2015.05.009

Hoffmann, B. S. (1974). *An examination of knowledge of results and step size in programmed instruction at high and low cognitive levels of objectives* (Publication No. 7501707) [Doctoral dissertation, Indiana University]. ProQuest Dissertations & Theses Global.

Hogg, P., Boyle, T., & Lawson, R. (1999). Comparative evaluation of a CORE based learning environment for nuclear medicine. *Journal of Educational Multimedia and Hypermedia, 8*, 457-473. doi:10.1111/j.1470-3327.2005.00614.x

Huang, J. C.-Y. (1995, February 8-10). *Effects of types of feedback on achievement and attitudes during computer-based cooperative condition* [Paper presentation]. Annual National Convention of the Association for Educational Communications and Technology, USA. Anaheim, CA, United States.

Hyman, C., & Tobias, S. (1981, October 21-23). *Feedback and prior achievement* [Paper presentation]. Annual Meeting of the North-eastern Educational Research Association, Ellenville, NY, United States.

Jacobs, P. I., & Kulkarni, S. (1963). A test of some assumptions underlying programed instruction. *Psychological Reports, 18,* 103-110. doi:10.1002/j.2333-8504.1963.tb00495.x

Kang, S. H. K., McDermott, K. B. & Roediger III, H. L., (2007) Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19*, 528-558. doi:10.1080/09541440601056620

Kauffman, D. F. (2004). Self-regulated learning in web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing, and motivational beliefs. *Journal of Educational Computing Research, 30*, 139-161. doi:10.2190%2FAX2D-Y9VM-V7PX-0TAD

Kealy, W. A. & Ritzhaupt, A. D. (2010). Assessment certitude as a feedback strategy for learners' constructed responses. *Journal of Educational Computing Research*, *43*, 25-45. doi:10.2190%2FEC.43.1.c

Krumboltz, J. D., & Weisman, R. G. (1962). The effect of intermittent confirmation in programed instruction. *Journal of Educational Psychology, 53*, 250-253. doi:10.1037/h0040448

Kulhavy, R. W., Yekovich, F. R., & Dyer, J. W. (1976). Feedback and response confidence. *Journal of Educational Psychology, 68*, 522-528. doi:10.1037/0022-0663.68.5.522

Kulhavy, R. W., Yekovich, F. R., & Dyer, J. W. (1979). Feedback and content review in programmed instruction. *Contemporary Educational Psychology, 4*, 91-98. doi:10.1016/0361-476X(79)90062-6

Lasoff, E. M. (1981). *The effects of feedback in both computer-assisted instruction and programmed instruction on achievement and attitude* (Publication No. 8121115) [Doctoral dissertation, University of Miami]. ProQuest Dissertations and Theses Global.

Lee, H. W., Lim, K. Y., & Grabowski, B. L. (2010). Improving self-regulation, learning strategy use, and achievement with metacognitive feedback. *Educational Technology Research and Development, 58*, 629-648. doi:10.1007/s11423-010-9153-6

Lhyle, K. G., & Kulhavy, R. W. (1987). Feedback processing and error correction. *Journal of Educational Psychology, 79*, 320-322. doi:10.1037/0022-0663.79.3.320

Lin, H. (2006). *The effect of questions and feedback used to complement static and animated visualization on tests measuring different educational objectives* (Publication No. 3318901) [Doctoral dissertation, The Pennsylvania State University]. ProQuest Dissertations & Theses Global.

Llorens, A. C., Cerdán, R., & Vidal-Abarca, E. (2014). Adaptive formative feedback to improve strategic search decisions in task-oriented reading. *Journal of Computer Assisted Learning, 30*, 233-251. doi:10.1111/jcal.12050

Llorens, A. C., Vidal-Abarca, E., & Cerdán, R. (2016). Formative feedback to transfer self-regulation of task-oriented reading strategies. *Journal of Computer Assisted Learning, 32*, 314-331. doi:10.1111/jcal.12134

Lublin, S. C. (1965). Reinforcement schedules, scholastic aptitude, autonomy need, and achievement in a programed course. *Journal of Educational Psychology, 56*, 295-302. doi:10.1037/h0022790

Martin, F., Klein, J. D., & Sullivan, H. (2007). The impact of instructional elements in computer-based instruction. *British Journal of Educational Technology, 38*, 623-636. doi:10.1111/j.1467-8535.2006.00670.x

Martin S. A., Azevedo R., Taub M., Mudrick N. V., Millar G. C., Grafsgaard J. F. (2016) Are there benefits of using multiple pedagogical agents to support and foster self-regulated learning in an intelligent tutoring system? In: A. Micarelli, J. Stamper, K. Panourgia (Eds.) *Intelligent tutoring systems: Lecture notes in computer science.* (Vol. 9684, pp. 273-279). doi:10.1007/978-3-319-39583-8_29

Merrill, D. M. (1965). Correction and review on successive parts in learning a hierarchical task. *Journal of Educational Psychology, 56*, 225-234. doi:10.1037/h0022471

Moore, J. W., & Smith, W. I. (1964). Role of knowledge of results in programed instruction. *Psychological Reports, 14*, 407-423. doi:10.2466%2Fpr0.1964.14.2.407

Morrison, G. R., Ross, S. M., Gopalakrishnan, M., & Casey, J. (1995). The effects of feedback and incentives on achievement in computer-based instruction. *Contemporary Educational Psychology, 20*, 32-50. doi:10.1006/ceps.1995.1002

Nelson, W. A. (1992, April 20-24). *Adaptive hypermedia instructional systems: Possibilities for learner modelling* [Paper presentation]. Annual Meeting of the American Educational Research Association. San Fransisco, CA, United States.

Nishikawa, S. (1988, January 14-19). *Effects of feedback strategies in computer assisted instruction and the influence of locus of control on the performance of junior high students* [Paper Presentation]. Annual Convention of the Association for Educational Communications and Technology. New Orleans, LA, United States.

Olson, G. H. (1971). *A multivariate examination of the effects of behavioral objectives, knowledge of results, and the assignment of grades on the facilitation of classroom learning* (Publication No. 7213552) [Doctoral dissertation, The Florida State University]. ProQuest Dissertations & Theses Global.

Peverly, S. T., & Wood, R. (2001). The effects of adjunct questions and feedback on improving the reading comprehension skills of learning-disabled adolescents. *Contemporary Educational Psychology, 26,* 25-43. doi:10.1006/ceps.1999.1025

Pridemore, D. R., & Klein, J. D. (1995). Control of practice and level of feedback in computer-based instruction. *Contemporary Educational Psychology, 20*, 444-450. doi:10.1006/ceps.1995.1030

Razagifard, P., Ghabelnezam, A., & Fard, V. F. (2011). The effect of computer-mediated feedback on second language reading comprehension. *International Journal on New Trends in Education and Their Implications, 2*(1), 90-98.

Ripple, R. E. (1963). Comparison of the effectiveness of a programed text with three other methods of presentation. *Psychological Reports, 12*, 227-237. doi:10.2466%2Fpr0.1963.12.1.227

Rothkopf, E. Z. (1966). Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. *American Educational Research Journal, 3*, 241-249. doi:10.3102%2F00028312003004241

Saunders, P. H. (1998). *The effectiveness of paraphrasing and feedback in computer-based instruction* (Publication No. 9905104) [Doctoral dissertation, The University of Memphis]. ProQuest Dissertations & Theses Global.

Sung, Y. T., Chang, K. E., & Huang, J. S. (2008). Improving children's reading comprehension and use of strategies through computer-based strategy training. *Computers in Human Behavior, 24,* 1552-1571. doi:10.1016/j.chb.2007.05.009

Tobias, S. (1987). Mandatory text review and interaction with student characteristics. *Journal of Educational Psychology, 79,* 154-161. doi:10.1037/0022-0663.79.2.154

Tobias, S. (1988). Teaching strategic text review by computer and interaction with student characteristics. *Computers in Human Behavior, 4*, 299-310. doi:10.1016/0747-5632(88)90002-7

Tobias, S., & Ingber, T. (1976). Achievement-treatment interactions in programmed instruction. *Journal of Educational Psychology, 68,* 43-47. doi:10.1037/0022-0663.68.1.43

Tsao, P. K. (1977). *The effects of error rate, knowledge of correct results and test anxiety in linear program* (Publication No. 7815385) [Doctoral dissertation, The University of Oklahoma]. ProQuest Dissertations & Theses Global.

Valdez, A. J. (2008). *Encouraging mindful feedback processing: Computer-based instruction in descriptive statistics* (Publication No. 3329482) [Doctoral dissertation, The University of New Mexico]. ProQuest Dissertations & Theses Global.

Wentling, T. L. (1973). Mastery versus nonmastery instruction with varying test item feedback treatments. *Journal of Educational Psychology, 65,* 50-58. doi:10.1037/h0034820

Appendix E

Overview of Studies Included in the Meta-Analysis, Including Moderator Variables and Effect Sizes

Authors	Year	Pub. Type	*N* fb	*N* cont	Age	Randomly assigned	Subject design	Nr. of sessions	Text type	Text design	Feedback Type	Feedback Timing	Means of providing feedback	Outcome measure	Effect size (Hedges'*g*)
Adams & Strickland	2012	J	17	17	Stu.	N	B	4	Inf.	Text only	KCR	After	Computer	MC-questions	0.00
Anderson et al.	1971A Exp. 1	J	21	21	Stu.	Y	B	1	Inf.	Text + pic.	KCR	During	Computer	Open + MC-questions	1.10
Anderson et al.	1971B Exp. 2	J	24	26	Stu.	Y	B	1	Inf.	Text + pic.	KCR	During	Computer	Open + MC-questions	0.60
Anderson et al.	1972	J	24#	24#	Stu.	Y	B	1	Inf.	Text only	KCR	During	Computer	Open + MC-questions	0.41
Azevedo et al.	2012	CP	23#	23#	Stu.	Y	B	1	Inf.	Text + pic.	EF	During	Computer	MC-questions	1.05
Butler & Roediger III	2008	J	26	26	Stu.	Y	W	1	Nar.	Text only	KCR	After	Computer	Open questions	0.52
Butler et al.	2013A	J	20	10*	Stu.	Y	B	1	Inf.	Text only	EF	After	Computer	Open questions	1.22
Butler et al.	2013B	J	20	10*	Stu.	Y	B	1	Inf.	Text only	KCR	After	Computer	Open questions	0.62
Chanond	1988	D	60	60	Stu.	Y	B	1	Inf.	Text + pic.	EF	During	Computer	MC-questions	0.77
														Del.: MC-questions	0.26
Chen et al.	2011A	J	19	19	Stu.	Y	B	1	Inf.	Text + video	KOR	During	Computer	MC-questions	0.05
Chen et al.	2011B	J	20	20	Stu.	Y	B	1	Inf.	Text + video	KOR	During	Computer	MC-questions	0.67
Clariana & Koul	2006	J	17	16	Sec.	Y	B	1	Inf.	Text only	KCR	After	Non-com.	MC-questions	0.19

4

Authors	Year	Pub. Type	*N* fb	*N* cont	Age	Randomly assigned	Subject design	Nr. of sessions	Text type	Text design	Feedback Type	Feedback Timing	Means of providing feedback	Outcome measure	Effect size (Hedges'*g*)
Clariana et al.	1992	CP	10	10	Sec.	Y	B	1	Inf.	Text only	KCR	-	Computer	MC-questions	2.40
														Del.: MC-questions	2.30
Clark & Dwyer	1998	J	37	47	Stu.	Y	B	1	Inf.	Text + pic.	EF	After	Computer	MC-questions	0.10
Elliot	1986	D	20	22	Sec.	Y	B	1	Inf.	Text only	EF	During	Computer	MC-questions	0.41
Farragher	1980A	D	30	10*	Stu.	Y	B	1	Inf.	Text only	KOR	During	Non-com.	MC-questions	-0.21
														Del.: MC-questions	-0.32
Farragher	1980B	D	30	10*	Stu.	Y	B	1	Inf.	Text only	EF	During	Non-com.	MC-questions	-0.18
														Del.: MC-questions	-0.24
Farragher	1980C	D	30	10*	Stu.	Y	B	1	Inf.	Tex only	EF	During	Non-com.	MC-questions	0.42
														Del.: MC-questions	0.28
Farragher & Yore	1997A	J	25	13*	Sec.	Y	B	1	Inf.	Text + pic.	EF	During	Non-com.	MC-questions	0.09
														Del.: MC-questions	-0.07
Farragher & Yore	1997B	J	25	12*	Sec.	Y	B	1	Inf.	Text + pic.	EF	During	Non-com.	MC-questions	0.09
														Del.: MC-questions	0.06
Fazio et al.	2010	J	48	48	Stu.	Y	W	1	Inf.	Text only	KCR	After	Computer	Open questions	1.21
Feldhusen & Birt	1962A	J	30	10*	Stu.	Y	B	1	Inf.	Text only	KCR	During	Computer	Questions (unkn.)	0.35
Feldhusen & Birt	1962B	J	30	10*	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	Questions (unkn.)	0.23

Authors	Year	Pub. Type	*N* fb	*N* cont	Age	Randomly assigned	Subject design	Nr. of sessions	Text type	Text design	Feedback Type	Feedback Timing	Means of providing feedback	Outcome measure	Effect size (Hedges'*g*)
Feldhusen & Birt	1962C	J	30	10*	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	Questions (unkn.)	0.34
Feng & Reigeluth	1983	Rep.	12	10	Pr.	N	B	1	Inf.	Text only	EF	During	Non-com.	MC-questions	1.31
Fernald & Jordon	1991A	J	18	18	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	MC-questions	0.11
Fernald & Jordon	1991B	J	18	16	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	MC-questions	0.05
Fernald & Jordon	1991C	J	16	15	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	MC-questions	-0.45
Franzke et al.	2005	J	52	59	Sec.	Y	B	8	Both	Text only	KOR	After	Computer	Questions (unkn.)	0.16
Goldbeck & Campbell	1962	J	16#	16#	Sec.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	Questions (unkn.)	-0.27
														Del.: Questions (unkn.)	-0.65
Golke et al.	2009	J	100	98	Sec.	Y	B	1	Both	Text only	KOR	After	Computer	MC-questions	-0.12
Golke et al.	2015A Exp. 1	J	108	28*	Sec.	Y	B	1	Both	Text only	EF	During	Computer	MC-questions	-0.18
														Del.: MC-questions	-0.12
Golke et al.	2015B Exp. 1	J	111	27*	Sec.	Y	B	1	Both	Text only	KOR	During	Computer	MC-questions	-0.11
														Del.: MC-questions	-0.01

Authors	Year	Pub. Type	*N* fb	*N* cont	Age	Randomly assigned	Subject design	Nr. of sessions	Text type	Text design	Feedback Type	Feedback Timing	Means of providing feedback	Outcome measure	Effect size (Hedges'*g*)
Golke et al.	2015C Exp. 1	J	113	27*	Sec.	Y	B	1	Both	Text only	EF	During	Computer	MC-questions	-0.22
														Del.: MC-questions	-0.09
Golke et al.	2015D Exp. 1	J	120	27*	Sec.	Y	B	1	Both	Text only	EF	During	Computer	MC-questions	0.11
														Del.: MC-questions	-0.02
Golke et al.	2015E Exp. 2	J	61	30*	Sec.	Y	B	1	Both	Text only	EF	During	Computer	MC-questions	0.37
Golke et al.	2015F Exp. 2	J	63	30*	Sec.	Y	B	1	Both	Tet only	EF	During	Computer	MC-questions	0.05
Hoffmann	1974	D	16#	16#	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	Open questions	-0.10
Hogg et al.	1999	J	8	10	Stu.	N	B	-	Inf.	Text + video	KOR	After	Computer	Open + MC-questions	0.81
Huang	1995A	CP	40	20*	Stu.	Y	B	2	Inf.	-	KCR	During	Computer	Del.: MC-questions	1.93
Huang	1995B	CP	40	20*	Stu.	Y	B	2	Inf.	-	EF	During	Computer	Del.: MC-questions	2.41[w]
Hyman & Tobias	1981	CP	46	42	Pr.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	Open questions	-0.19
Jacobs & Kulkarni	1966A	J	15	12	Sec.	Y	B	-	Inf.	Text + pic.	KCR	During	Non-com.	Open + MC-questions	-0.91
Jacobs & Kulkarni	1966B	J	19	9	Sec.	Y	B	-	Inf.	Text + pic.	KCR	During	Non-com.	Open + MC-questions	-0.39
Kang et al.	2007A	J	48	24*	Stu.	Y	W	1	Inf.	Text only	KCR	After	Computer	Open + MC-questions	0.85

Authors	Year	Pub. Type	*N* fb	*N* cont	Age	Randomly assigned	Subject design	Nr. of sessions	Text type	Text design	Feedback Type	Feedback Timing	Means of providing feedback	Outcome measure	Effect size (Hedges'*g*)
Kang et al.	2007B	J	48	24*	Stu.	Y	W	1	Inf.	Text only	KCR	After	Computer	Open + MC-questions	1.23
Kauffman	2004A	J	15#	15#	Stu.	Y	B	1	Inf.	Text only	KOR	After	Computer	Open + MC-questions	0.41
Kauffman	2004B	J	15#	15#	Stu.	Y	B	1	Inf.	Text only	KOR	After	Computer	Open + MC-questions	0.53
Kauffman	2004C	J	15#	15#	Stu.	Y	B	1	Inf.	Text only	KOR	After	Computer	Open + MC-questions	0.05
Kauffman	2004D	J	15#	15#	Stu.	Y	B	1	Inf.	Text only	KOR	After	Computer	Open + MC-questions	-0.03
Kealy & Ritzhaupt	2010	J	21	19	Stu.	Y	B	1	Nar.	Text only	KCR	After	Computer	Open questions + recall	0.32
Krumboltz & Weisman	1962	J	20	25	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	Open questions	0.18
Kulhavy et al.	1976	J	30	30	Stu.	Y	B	1	Inf.	Text + pic.	KOR	During	Non-com.	MC-questions	0.51
														Del.: MC-questions	0.55
Kulhavy et al.	1979	J	60	60	Stu.	Y	B	1	Inf.	Text + pic.	KOR	During	Non-com.	MC-questions	0.70
Lasoff	1981A	D	12	12	Stu.	Y	B	5	Inf.	Text only	EF	During	Computer	MC-questions	-0.32
Lasoff	1981B	D	10	10	Stu.	Y	B	5	Inf.	Text only	EF	During	Non-com.	MC-questions	0.00
Lee et al.	2010	J	74#	74#	Stu.	Y	B	1	Inf.	Text + pic.	KOR	During	Computer	MC-questions	0.54

Authors	Year	Pub. Type	*N* fb	*N* cont	Age	Randomly assigned	Subject design	Nr. of sessions	Text type	Text design	Feedback Type	Feedback Timing	Means of providing feedback	Outcome measure	Effect size (Hedges'*g*)
Lhyle & Kulhavy	1987	J	20	20	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	MC-questions	1.26
Lin	2006	D	194	194	Stu.	Y	B	1	Inf.	Text + video	EF	During	Computer	Open + MC-questions	0.10
Llorens et al.	2014A	J	30	14*	Sec	Y	B	1	Inf.	Text + pic.	EF	During	Computer	MC-questions	0.67
Llorens et al.	2014B	J	34	14*	Sec.	Y	B	1	Inf.	Text + pic.	KCR	During	Computer	MC-questions	0.35
Llorens et al.	2016A Exp. 1	J	51	25*	Sec.	Y	B	1	Inf.	Text only	EF	During	Computer	MC-questions	0.10
Llorens et al.	2016B Exp. 1	J	41	25*	Sec.	Y	B	1	Inf.	Text only	EF	During	Computer	MC-questions	0.16
Llorens et al.	2016C Exp. 2	J	38	37	Sec.	Y	B	1	Inf.	Text only	EF	During	Computer	MC-questions	0.05
Lublin	1965	J	55#	55#	Stu.	Y	B	12	Inf.	Text + pic.	KCR	During	Non-com.	Open questions	-0.55
Martin et al.	2007	J	43#	43#	Stu.	Y	B	1	Inf.	Text + pic.	KOR	During	Computer	MC-questions	1.11
Martin et al.	2016	CP	82#	83#	Stu.	Y	B	2	Inf.	Text + pic.	EF	During	Computer	MC-questions	0.41
Merrill	1965	J	14	12	Stu.	Y	B	-	Inf.	Text only	EF	During	Computer	MC-questions	2.03
Moore & Smith	1964A	J	22	22	Stu.	Y	B	12	Inf.	Text only	KCR	During	Computer	Open + MC-questions	0.47

Authors	Year	Pub. Type	*N* fb	*N* cont	Age	Randomly assigned	Subject design	Nr. of sessions	Text type	Text design	Feedback Type	Feedback Timing	Means of providing feedback	Outcome measure	Effect size (Hedges'*g*)
Moore & Smith	1964B	J	22	22	Stu.	Y	B	12	Inf.	Text only	KCR	During	Computer	Open + MC-questions	0.92
Morrison et al.	1995A	J	51	24*	Stu.	Y	B	1	Inf.	Text only	KCR	During	Computer	MC-questions	0.24
Morrison et al.	1995B	J	49	24*	Stu.	Y	B	1	Inf.	Text only	KOR	During	Computer	MC-questions	-0.18
Nelson	1992	CP	12	12	Stu	Y	B	1	Inf.	Text only	KOR	During	Computer	Open questions	0.32
Nishikawa	1988	CP	10	13	Sec.	N	B	1	Inf.	Text + pic.	KCR	During	Computer	Questions (unkn.)	0.08
														Del.: Questions (unkn.)	0.01
Olson	1971A	D	13	12	Stu.	Y	B	4	Inf.	Text + pic.	KOR	After	Non-com.	MC-questions	0.40
														Del.: MC-questions	0.27
Olson	1971B	D	12	13	Stu.	Y	B	4	Inf.	Text + pic.	KOR	After	Non-com.	MC-questions	0.51
														Del.: MC-questions	0.24
Olson	1971C	D	13	13	Stu.	Y	B	4	Inf.	Text + pic.	KOR	After	Non-com.	MC-questions	0.09
														Del.: MC-questions	0.00
Olson	1971D	D	13	12	Stu.	Y	B	4	Inf.	Text + pic.	KOR	After	Non-com.	MC-questions	-0.05
														Del.: MC-questions	-0.13
Peverly & Wood	2001A	J	10	10	Sec.	Y	B	6	Nar.	Text only	KCR	During	Non-com.	Open questions	0.12
Peverly & Wood	2001B	J	10	10	Sec.	Y	B	6	Nar.	Text only	KCR	After	Non-com.	Open questions	0.68
Pridemore & Klein	1995A	J	35#	18#*	Sec.	Y	B	2	Inf.	Text + pic.	EF	During	Computer	Open + MC-questions	0.07

4

Authors	Year	Pub. Type	*N* fb	*N* cont	Age	Randomly assigned	Subject design	Nr. of sessions	Text type	Text design	Feedback Type	Feedback Timing	Means of providing feedback	Outcome measure	Effect size (Hedges'*g*)
Pridemore & Klein	1995B	J	35#	17#*	Sec.	Y	B	2	Inf.	Text + pic.	KCR	During	Computer	Open + MC-questions	-0.22
Pridemore & Klein	1995C	J	35#	18#*	Sec.	Y	B	2	Inf.	Text + pic.	EF	During	Computer	Open + MC-questions	0.06
Pridemore & Klein	1995D	J	35#	17#*	Sec.	Y	B	2	Inf.	Text + pic.	KCR	During	Computer	Open + MC-questions	-0.15
Razagifard et al.	2011	J	30	30	Stu.	Y	B	4	-	Text only	KOR + KCR	-	Computer	Questions (unkn.)	0.94
Ripple	1963	J	60	60	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	Open + MC-questions Del.: Open + MC-questions	0.01 0.21
Rothkopf	1966	J	21	21	Stu.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	Open questions	0.46
Saunders	1998A	D	17	17	Stu	Y	B	1	Inf.	Text only	KCR	After	Computer	Open + MC-questions	-0.14
Saunders	1998B	D	17	17	Stu	Y	B	1	Inf.	Text only	KCR	After	Computer	Open + MC-questions	-0.44
Sung et al.	2008A	J	31	35	Pr.	N	B	22	Both	Text only	EF	After	Non-com.	MC-questions	0.22
Sung et al.	2008B	J	34	30	Pr.	N	B	22	Both	Text only	EF	After	Non-com.	MC-questions	0.54
Tobias	1987A	J	37	19*	Sec.	Y	B	1	Inf.	Text only	KOR + KCR	During	Computer	Open questions	0.20
Tobias	1987B	J	37	18*	Sec.	Y	B	1	Inf.	Text only	KOR + KCR	During	Computer	Open questions	0.20

Authors	Year	Pub. Type	*N* fb	*N* cont	Age	Randomly assigned	Subject design	Nr. of sessions	Text type	Text design	Feedback Type	Feedback Timing	Means of providing feedback	Outcome measure	Effect size (Hedges'*g*)
Tobias	1988A	J	25	25	Sec.	Y	B	1	Inf.	Text only	KOR + KCR	During	Computer	Open questions	0.21
Tobias	1988B	J	20	23	Sec.	Y	B	1	Inf.	Text only	KOR + KCR	During	Computer	Open questions	0.46
Tobias & Ingber	1976	J	52#	52#	Pr.	Y	B	1	Inf.	Text only	KCR	During	Non-com.	Open questions	0.15
Tsao	1977A	D	20	20	Stu.	Y	B	1	Inf.	-	KCR	During	Non-com.	Open questions	0.19
Tsao	1977B	D	20	20	Stu.	Y	B	1	Inf.	-	KCR	During	Non-com.	Open questions	-0.11
Valdez	2008A	D	22#	11#*	Stu.	Y	B	1	Inf.	Text + pic.	EF	After	Computer	MC-questions	1.94
Valdez	2008B	D	22#	10#*	Stu.	Y	B	1	Inf.	Text + pic.	KCR	After	Computer	MC-questions	2.25
Valdez	2008C	D	22#	11#*	Stu.	Y	B	1	Inf.	Text + pic.	EF	After	Computer	MC-questions	2.41w
Valdez	2008D	D	22#	10#*	Stu.	Y	B	1	Inf.	Text + pic.	KCR	After	Computer	MC-questions	2.40w
Wentling	1973A	J	39#	19#	Sec.	N	B	20	Inf.	Text only	KOR	After	Non-com.	MC-questions	0.17
														Del.: MC-questions	-0.04
Wentling	1973B	J	39#	20#	Sec.	N	B	20	Inf.	Text only	KOR	After	Non-com.	MC-questions	-0.31
														Del.: MC-questions	-0.20

Note: Publication type: CP = conference proceeding, D = dissertation, J = journal article, Rep. = report. Age: Pr. = primary school students, Sec. = secondary school students, *N* fb = number of participants in the feedback condition. *N* cont = number of participants in the control condition. Stu. = university students. Randomly assigned: Y = yes, random assignment of participants to condition on an individual level, N = no random assignment or randomization at classroom/group level. Subject design: B = between-subject design, W = within-subject design. Text type: Inf. = informative text(s), Nar. = narrative text(s), Both = informative and narrative texts. Text design: Text + pic. = text accompanied by pictures, Text + video = text accompanied by animations or video content. FB Type: KOR = knowledge-of-response feedback, KCR = knowledge-of-correct-response feedback, EF = elaborate feedback. Means of providing feedback: non-com. = non-computer-delivered, Computer = computer-delivered. Outcome measure: Del. = delayed post-test, Questions (unkn.) = type of comprehension questions is unknown. # = number of participants estimated based on the total sample size and the number of conditions in a study. * Number of participants in the control group was split among multiple contrasts. ^{w} = winsorized effect size