Maastricht University

The Effects of Summarization and Factual Retrieval Practice on Text Comprehension and Text Retention in Elementary Education

Citation for published version (APA):

Ophuis-Cox, F. H. A., Rozendal, L., Catrysse, L., Joosten-ten Brinke, D., & Camp, G. (2023). The Effects of Summarization and Factual Retrieval Practice on Text Comprehension and Text Retention in Elementary Education. *Journal of Experimental Psychology-Applied*, *30*(2), 258-267. https://doi.org/10.1037/xap0000507

Document status and date:
Published: 21/12/2023

DOI:
10.1037/xap0000507

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Please check the document version of this publication:

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

Link to publication

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

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

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

www.umlib.nl/taverne-license

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

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 17 Nov. 2026

Journal of Experimental Psychology: Applied

The Effects of Summarization and Factual Retrieval Practice on Text Comprehension and Text Retention in Elementary Education

F. H. A. Ophuis-Cox, L. Rozendal, L. Catrysse, D. Joosten-ten Brinke, and G. Camp

Online First Publication, December 21, 2023. https://dx.doi.org/10.1037/xap0000507

CITATION

Ophuis-Cox, F. H. A., Rozendal, L., Catrysse, L., Joosten-ten Brinke, D., & Camp, G. (2023, December 21). The Effects of Summarization and Factual Retrieval Practice on Text Comprehension and Text Retention in Elementary Education. *Journal of Experimental Psychology: Applied*. Advance online publication. https://dx.doi.org/10.1037/xap0000507

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Applied

© 2023 American Psychological Association
ISSN: 1076-898X

https://doi.org/10.1037/xap0000507

The Effects of Summarization and Factual Retrieval Practice on Text Comprehension and Text Retention in Elementary Education

F. H. A. Ophuis-Cox[1], L. Rozendal[1], L. Catrysse[1], D. Joosten-ten Brinke[2], and G. Camp[1]

[1] Department of Educational Sciences, Open University of the Netherlands

[2] Department of Health, Medicine and Life Sciences, Maastricht University

When reading a text in school, the goal is both text comprehension and text retention. We examined the effects of the learning strategies summarization and factual retrieval practice on third- and fourth-grade pupils' text comprehension and retention of factual knowledge from a text, using restudy as a control condition. The experiment was conducted in an authentic classroom setting, with teachers executing the experiment using original course materials. In 2016, 57 regular third- and fourth-grade pupils ($M = 9.04$ years old) read three different texts, and each applied three different learning strategies (summarization, retrieval practice and restudy, which were counterbalanced across texts) in subsequent practice sessions. After a 2-week delay, a final test was administered. The learning strategy summarization had a larger positive effect on text comprehension than factual retrieval practice, but had a similar effect compared to restudy. The learning strategy factual retrieval practice had a larger positive effect on text retention than both summarization and restudy. Implications for educational practice are discussed.

Public Significance Statement

We show that, compared to factual retrieval practice, summarization and restudy have a larger positive effect on third- and fourth-grade pupils' text comprehension in an authentic elementary classroom setting. The summarization effects were comparable to the more time-efficient learning strategy restudy, implying that for summarization to be beneficial at this age, a summary training is necessary to increase pupils' summary quality. We show that, compared to summarization and restudy, factual retrieval practice is the most beneficial learning strategy for the retention of factual knowledge from a text in third- and fourth-grade pupils and could therefore be implemented in elementary school curricula for this purpose.

Keywords: retrieval practice, summarization, memory, reading comprehension, elementary education

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

In elementary education, important foundations are laid when it comes to reading and understanding text material. At the age of 8–14, a transition from learning basic reading skills to the "reading-for-learning" stage takes place (Chall, 1983). Educators hope to achieve that, when studying a text, elementary school pupils will understand this text at a deep level (text comprehension) and will store its detailed facts (factual text retention), such as definitions, names and years in long-term memory (Kendeou et al., 2014; Organization for Economic Cooperation and Development [OECD], 1999). To reach these two different goals, the learning strategy summarization is often advised in learning methods and instructed by elementary school teachers (National Reading Panel et al., 2000; New South Wales Centre for Effective Reading, 2022).

Overall, literature supports the positive effect of writing a summary on both text retention (e.g., Westby et al., 2010; Wormeli, 2004) and text

F. H. A. Ophuis-Cox https://orcid.org/0000-0001-9985-3019

L. Rozendal https://orcid.org/0009-0009-0718-9612

L. Catrysse https://orcid.org/0000-0003-1537-1632

D. Joosten-ten Brinke https://orcid.org/0000-0001-6161-7117

G. Camp https://orcid.org/0000-0002-5272-150X

This publication is part of the project "The implementation of Retrieval Practice to Improve Language and Math Skills in Primary Education" with project No. 023.019.001 (awarded to F. H. A. Ophuis-Cox) of the research program "Doctoral Grant for Teachers" which is (partly) financed by the Dutch Research Council (NWO). As an employer of F. H. A. Ophuis-Cox, MSc, PhD candidate, the Board of Stichting Katholiek en Protestants-Christelijk Onderwijs Eindhoven provided financial support by facilitating time for this research. The authors thank the directors, teachers, and children of the primary schools for their cooperation in this research.

The authors have no conflicts of interest to disclose.

F. H. A. Ophuis-Cox played a lead role in formal analysis, funding acquisition, and writing–original draft and an equal role in validation and visualization. L. Rozendal played a lead role in data curation, investigation, project administration, and resources, a supporting role in writing–review and editing, and an equal role in conceptualization, methodology, and validation. L. Catrysse played a supporting role in formal analysis, supervision, and writing–original draft and an equal role in writing–review and editing. D. Joosten-ten Brinke played a supporting role in supervision and writing–review and editing. G. Camp played a lead role in methodology, supervision, and writing–review and editing, a supporting role in data curation, formal analysis, funding acquisition, project administration, validation, visualization, and writing–original draft and an equal role in conceptualization.

Correspondence concerning this article should be addressed to F. H. A. Ophuis-Cox, Department of Educational Sciences, Open University of the Netherlands, Valkenburgerweg 177, 6419 AT Heerlen, The Netherlands. Email: fieke.ophuis-cox@ou.nl

comprehension in different age groups (e.g., Nurhayati & Fitriana, 2018; Oded & Walters, 2001; Pakzadian & Rasekh, 2012; Rinehart et al., 1986; Taylor, 1982 and see Stevens et al., 2018, for a review). However, despite positive results, these experiments often include extensive summarization training programs, that are not commonly used in authentic classroom settings. Therefore, these studies do not represent authentic classroom research. Also, in most studies, the effect of summarization on text comprehension and text retention is compared to a restudy control condition. Despite the importance of this comparison, it would be interesting to compare summarization to other, more promising learning strategies, such as retrieval practice (i.e., retrieving to-be-learned information from memory).

Especially in the case of retaining factual knowledge, retrieval practice shows robust positive effects in different domains and for all age groups (for reviews, see Adesope et al., 2017; Agarwal et al., 2021; Brod, 2021; Dunlosky et al., 2013). Although positive effects have been found for the retention of text materials in adolescents and adults (e.g., Agarwal, 2019; McDermott et al., 2014; Rowley & McCrudden, 2020), elementary school classroom research is scarce and limited to the domain of vocabulary learning (Goossens et al., 2014) and spelling (Jones et al., 2016; Karpicke et al., 2016). Because effects of learning strategies can differ per age group, and elementary school pupils are known to struggle when learning strategies have to be applied effectively and efficiently (Brod, 2021), it is unclear if the positive effect on the retention of text material, found in older students, holds in elementary school pupils. Additionally, when it comes to the effect of retrieval practice on text comprehension, research is still in its infancy and is limited to middle school students and undergraduates, showing mixed results (e.g., Agarwal, 2019; Butler, 2010).

Thus, a number of studies have examined the effects of the learning strategies summarization and retrieval practice on text comprehension and text retention. However, these experiments either did not represent a classroom situation (using extensive summarization training programs), were conducted with adolescents or adults (instead of elementary school pupils), or used a restudy control condition (instead of comparing two promising learning strategies). Therefore, we wanted to conduct a study that compared the effects of summarization and retrieval practice on the text comprehension and text retention of elementary school pupils in an authentic classroom setting.

Text Retention and Text Comprehension

To remember and to understand a text, are two different things (Kintsch, 1994). First, remembering a text means that the reader is able to reproduce information that was described in the text (Kintsch, 1994). Text retention is referred to as the retention of factual knowledge that can be found verbatim within the text, such as names, dates, examples, and definitions (Kendeou et al., 2014; OECD, 1999). Second, understanding a text implies that a reader is able to use the information that is provided by the text in other ways than just for reproduction (Kintsch, 1994). To understand a text in a way that the learner is able to answer questions and verify statements about it or paraphrase it, several levels of comprehension have to be successfully completed (Kintsch, 1988, 1994; Van Dijk & Kintsch, 1983). Kintsch tried to capture these levels in his prominent construction-integration model of text comprehension (Kintsch, 1988; McNamara & Magliano, 2009). In this theory, Kintsch (1988, 1994) distinguished the text-base model and the situation model. The text-base model includes text information and the relation between propositions in the text. The situation model includes the integration of the text's information with the readers' background knowledge (Kintsch, 1994; McNamara & Magliano, 2009; Meyer & Ray, 2011; Van Dijk & Kintsch, 1983). Within these two models, different levels of understanding occur. There is a more superficial state of understanding as in reproducing a relation between propositions in a text (the text-based level) and a deeper level of understanding in which such relations are linked to the prior knowledge and lead to a level of understanding beyond the information that was made explicit in the text itself (Kintsch, 1994).

Summarization

There are two ways to write a summary. If writing a summary encourages a learner to integrate new knowledge into prior knowledge, thereby facilitating deeper understanding and prompting a learner to construct something meaningful that goes beyond the information that is given in the text, it counts as a generative learning strategy (Brod, 2021). If writing a summary encourages learners to execute the procedure of paraphrasing and condensing the text, in which the main thoughts of a text are transformed to a shorter version, it is known as the rule-based summarizing strategy (Brown et al., 1981). This second method of summarization is suitable for elementary school students in the lower grades of elementary education, given their stage of reading and writing development (CED-groep, Centrum Educatieve Dienstverlening, 2022; Chall, 1983; Christie, 2010). Following this strategy, first, the reader deletes all trivial material unnecessary to understanding. Second, redundant material is deleted. Third, superordinate terms are used for lists or examples, and finally a topic sentence is selected or invented if missing in the text. Because of the focus on the main thoughts, concepts and interrelationships of a text, it can be expected that writing a summary leads to greater text comprehension (Marzano et al., 2001; Oded & Walters, 2001). This statement is supported by several experimental studies showing positive effects of summarization on text comprehension in adolescents (e.g., Nurhayati & Fitriana, 2018; Oded & Walters, 2001; Pakzadian & Rasekh, 2012), and also in elementary school pupils (e.g., Rinehart et al., 1986; Taylor, 1982).

As for text retention, it is argued that writing a summary is effective for memorizing information from a text (e.g., Westby et al., 2010; Wormeli, 2004). However, it is questionable whether this effect remains when the retention of *factual* information is the goal, because, when writing a summary a reader focusses on capturing the main thoughts of a text instead of memorizing the facts and details (Brown et al., 1983; Hagaman et al., 2016; Hidi & Anderson, 1986; Oded & Walters, 2001). It is plausible that, during this process, factual knowledge, such as names and definitions, is more likely to be forgotten. Therefore, it might be expected that the positive effect of summarization remains absent when the retention of factual knowledge is the goal. This hypothesis is supported by an experiment with undergraduates that found no advantage of the learning strategy summarization on factual text retention compared to the learning strategy restudy (Spirgel & Delaney, 2016). Additionally, Mok and Chan (2016) compared the learning strategy summarization with the learning strategy retrieval practice, showing an advantage for retrieval practice on the text retention of 12- to 14-year-old students with low

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

test anxiety. Hence, for the retention of factual knowledge from a text, it is questionable whether the often-used learning strategy summarization is beneficial compared to other learning strategies, such as restudy or factual retrieval practice, when tested in an authentic elementary school setting.

Retrieval Practice

Positive effects of retrieval practice on the retention of factual knowledge from text materials have been found in middle school students (e.g., McDaniel et al., 2013; McDermott et al., 2014; Rowley & McCrudden, 2020), college students, and undergraduates (e.g., Agarwal, 2019; Butler, 2010; Chan et al., 2006; McDaniel et al., 2009; Roediger & Karpicke, 2006). The effect of retrieval practice on factual text retention has not yet been investigated in an elementary school classroom setting (see Agarwal et al., 2021, for a review). However, because of positive effects found in other domains of authentic elementary school settings (e.g., Goossens et al., 2014; Jones et al., 2016; Karpicke et al., 2016), it might be expected that these positive effects generalize to the domain of learning from text materials.

As for text comprehension, mixed results have been found for retrieval practice in experiments with adolescents that investigated possible transfer effects of factual retrieval practice questions (e.g., Agarwal, 2019; Butler, 2010; Chan et al., 2006; Karpicke & Blunt, 2011; McDaniel et al., 2013). For example, undergraduates that practiced with factual retrieval questions to retain knowledge from text material, scored better on a 1-week delayed test with new inferential questions about that text (Butler, 2010). It has to be mentioned that, in this experiment, the factual retrieval questions were highly related to the new inferential questions. Other experiments contradicted this effect (e.g., Agarwal, 2019; McDaniel et al., 2013). For instance, practicing with factual definition questions did not improve the answers of middle school students on questions where these definitions had to be applied (McDaniel et al., 2013). Also, Agarwal (2019, Experiments 1 and 2) found that factual quizzes for college students had no positive effect on higher order questions about that same text, even though the questions were related. However, in Agarwal's (2019) third experiment, a positive effect of mixed quizzes (quizzes that contained both factual and higher order questions), compared to purely higher order quizzes, was found on final test higher order questions for sixth-grade students. From these findings, it can be argued that the factual questions did have some kind of transfer effect. Unfortunately, this experiment did not include a condition in which only factual questions were used.

It is plausible to expect that practicing with pure factual retention questions, not paying any attention to the main thoughts of a text, will not enhance text comprehension. However, it remains unclear if recalling factual knowledge from a text contributes to understanding that knowledge at a level that inferences can be made to enhance text comprehension (Nguyen & McDaniel, 2016).

As described above, when learning from text materials, text comprehension and text retention are two learning goals to keep in mind. It is, however, unclear which learning strategy is best used to reach each of these goals in an authentic elementary school setting. Therefore, we examined the following research question: What is the effect of the learning strategies summarization, factual retrieval practice, and restudy (control condition) on third- and fourth-grade pupils' text comprehension and text retention?

Due to the focus of summarization on capturing the main thoughts of a text, which promotes a deeper understanding (Marzano et al., 2001; Oded & Walters, 2001), and the emphasis of factual retrieval practice on retaining factual knowledge derived from a text (Adesope et al., 2017; Agarwal et al., 2021; Dunlosky et al., 2013), the research question was accompanied by the following two hypotheses: (1) The learning strategy summarization has a larger effect on third- and fourth-grade pupils' text comprehension than the learning strategies retrieval practice and restudy, and (2) the learning strategy retrieval practice has a larger effect on third- and fourth-grade pupils' text retention than the learning strategies summarization and restudy.

Method

Design

In this experiment, a within-subjects design was used. After an initial learning phase, participants practiced the content of three different texts using three different learning strategies. To reduce chances of bias caused by a within-design (e.g., carryover effects), learning strategies were counterbalanced across texts (see Table 1). The first independent variable "learning strategy" was manipulated within-subjects and had three levels: retrieval practice, summarization, and restudy (control condition). The second independent variable "test type" had two levels: comprehension questions and factual retention questions. The dependent variable was the test score on the comprehension questions and the factual retention questions in the final test. This final test was administered 2 weeks after the last practice session. Additionally, during practice sessions, the performance on retrieval practice tests and the quality of summaries were measured. For a schematic overview of the design see Table 2.

Participants

Based on a power calculation (G*Power, Version 3.1.9.7) with an effect size of Cohen's f 0.25 (Cohen, 1988), and a power of 95%, 45 participants had to be recruited. The effect size was based on similar effect sizes found for retrieval practice compared to restudy (e.g.,

Table 1
Counterbalancing of Strategies Across Texts

Text	First group	Second group	Third group
Text 1: flood disaster	Restudy	Summarization	Retrieval practice
Text 2: self-driving trucks	Summarization	Retrieval practice	Restudy
Text 3: 3D printers	Retrieval practice	Restudy	Summarization

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2
Distribution of Instruction Lesson (ILS), Practice Session (PS), and Test Sessions

Date	Week 1 Thursday	Week 2 Monday	Week 2 Thursday	Week 3 Monday	Week 3 Thursday	Week 4 Monday	Week 4 Thursday	Week 5 Thursday	Week 6 Thursday
Activity	Text1, ILS and PS 1	Text 1, PS 2	Text 1, PS 3 Text 2, ILS 1 and PS 1	Text 2, PS 2	Text 2, PS 3 Text 3, ILS 1 and PS 1	Text 3, PS 2	Text 3, PS 3 Test Session 1, Text 1	Test Session 2, Text 2	Test Session 3, Text 3

Rowland, 2014), summarization compared to restudy (e.g., Stevens et al., 2018), and summarization compared to retrieval practice and restudy (Mok & Chan, 2016). Therefore, the parents of 76 third- and fourth-grade pupils of a regular Dutch elementary school were approached. The school had a school-weight of "zero" (Mulder & Meijnen, 2013), which implies that, based on the socioeconomic status of parents, no educational disadvantage is to be expected. By means of an informed consent, parents were asked to give permission for the pupil's participation. Only the data of pupils that participated in all sessions were analyzed. Two parents did not give consent, and 17 pupils were absent during one or more sessions. This resulted in a sample size of 57 pupils, 33 boys (57,9%), and 24 girls (42,1%), with an average age of 9.04 ($SD = 0.63$).

Materials

Texts

The three texts used in this experiment were selected from "Nieuwsbegrip", a Dutch learning method for reading comprehension (CED-groep, Centrum Educatieve Dienstverlening, 2022). Since 2006, this is a widely used method in Dutch elementary schools. We selected three A-level texts from this method, representing an A2 CEF language level (Common European Framework of Reference for Languages) and a 156–186 CLIB-score (Cito LeesIndex voor het basisonderwijs-score: a score to represent the degree of difficulty regarding text understanding), which is suitable for third- and fourth-grade pupils (CED-groep, Centrum Educatieve Dienstverlening, 2022). Text 1 "Flood disaster" consists of 353 words, Text 2 "Self-driving trucks" of 374 words, and Text 3 "The rise of the 3D printer" of 372 words. All three texts were divided into five paragraphs with a heading. To prevent pupils from having studied the texts before, the texts were selected from the method's archive.

Summaries

For each text, a high-quality summary (expert summary) was made as prescribed by the comprehension method, based on the rule-based summarization strategy (Brown et al., 1981; Marzano et al., 2001). Each text consisted of five paragraphs. The expert summary consisted of the five headings of each paragraph containing two or three main thoughts (sentences) per paragraph, selected by the second researcher. The thoughts were directly converted from the text by paraphrasing or condensing the text. From pupils at the age of 8–10 years old, it can be expected that their reading and writing skills have reached the stage in which they are able to produce this form of summary (Chall, 1983; Christie, 2010).

Tests

For each of the three texts, the second author developed a written final test with 15 questions (10 factual retention questions and five comprehension questions). The goal of these questions was to measure the effect of the learning strategies on text-based comprehension and text-based retention. To measure text-based comprehension, bridging inference and sorting questions were used (see Thiede et al., 2012), to which the answer could be found in the text. Bridging inference questions are questions that deal with specific information that addresses the underlying text structure. Two examples of inference questions from the test about the flood disaster were: "Can you

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

describe what the flooded land looked like?" and "Can you name a difference and a similarity between a dike and a storm surge barrier?" Sorting questions are questions that measure the relations that students have found in the text and have integrated into their long-term memory (McNamara et al., 1996). For these sorting questions, the pupils needed to determine to which paragraph a statement belonged. For example, the "Flood disaster" text contained a paragraph with the heading "Prevention". In this paragraph, the reader received information about preventing flood disasters from happening. The corresponding sorting questing gave the following text-based statements, which participants had to place back in the right category: "Such a disaster should never happen again. The dikes along the sea and the rivers were strengthened. People also built dams and storm surge barriers." Participants could choose from the categories: Introduction, Storm, Help, Prevent, and Future. These categories represented the headings of the five paragraphs from the text. In this case, the category "Prevent" was the right answer.

To measure factual text retention, it was made sure that the answer to a question could be found verbatim in the text. The 10 factual retention questions referred to objective knowledge, such as names, dates, examples, and definitions, and were short-answer questions. Two examples of factual retention questions about the flood disaster were: "In what year did the flood disaster take place?" and "What is a storm surge barrier?" These 10 retention questions were used for both the retrieval practice sessions and the final retention test.

Procedure

Prior to the study, a pilot was conducted with a group of 10 pupils, who did not participate in the experiment. The second author observed and analyzed how much time students needed to read the texts and complete the assignments. The scheduled times, based on the times indicated in the learning method (CED-groep, Centrum Educatieve Dienstverlening, 2022), were sufficient (10 min for the restudy sessions, 15 min for the retrieval practice sessions, and 20 min for the summarization sessions), therefore, no adjustments were made. Additionally, this group of students read the text "Flood disaster" and practiced with the retrieval practice strategy. As a result of this pilot, several retention and comprehension questions were adapted.

The experiment was conducted in three existing third- and fourth-grade classrooms of elementary education. Three teachers in three different classrooms read the same text out loud. Afterward, in each classroom, pupils performed three practice sessions using a different learning strategy for each text (summarization, retrieval practice, and restudy). Three practice sessions were chosen, because when learning from text materials distributed practice is advised (see Cepeda et al., 2006, for a review). In the summarization and retrieval practice conditions, feedback was given to mimic the authentic way of teaching, and because of its moderating effect on learning (Kornell et al., 2009; Lipko-Speed et al., 2014; Marsh et al., 2012; Pashler et al., 2005). To ensure that the same procedure was used during the practice sessions across groups, each group received scripted instructions from their teacher, under supervision of the second author.

In the summarization condition, the teacher provided an instruction lesson (10 min) on how to write a summary, derived from the reading comprehension method's manual (CED-groep, Centrum Educatieve Dienstverlening, 2022). A sample text ("Airport Schiphol has existed for 100 years") and the digital whiteboard were used to model an expert summary. After this modeling session, the pupils received the experimental text and read it for 10 min. Then, they independently started making a summary of the experimental text, according to the rule-based summarization strategy (Brown et al., 1981; Marzano et al., 2001). They underlined important words and sentences before writing a summary on a blank sheet. During this process the text could be consulted. After 20 min, the teacher provided classroom feedback (10 min), by showing the text's expert summary on the digital whiteboard, so pupils could compare their own selected sentences to the expert idea units. To limit the difference in time on task and to mimic the spaced practice in the retrieval practice group, two additional practice sessions were also given in the summary group. During these two practice sessions the pupils once more restudied the text (10 min) and made a new summary on a blank sheet (20 min). At the end of these sessions, again classroom feedback was provided (10 min).

In the retrieval practice condition, pupils restudied the text for 10 min, and thereafter received 10 retention questions on factual knowledge, that were equal to the 10 final test retention questions. They had 15 min to answer these questions. This test was followed by a 10 min feedback session for the group as a whole, in which the teacher used the digital whiteboard to provide the right answers to the questions. The students checked their work and had the opportunity to ask questions. During the second and third practice sessions, the text was no longer consulted. The pupils only received the same 10 questions as in practice session one, followed by a feedback session.

In the restudy condition (control condition), pupils independently reread and restudied the text for 10 min. A time frame of 10 min was chosen, because within this time frame pupils were able to at least reread the text twice. No additional time was given, as research shows that additional reading sessions show diminishing efficiency (Rothkopf, 1968). Also, giving additional time would not represent an authentic classroom situation. For this classroom design, the 10 min time frame was expected to generate optimal effects for the rereading condition.

Two weeks after the last practice session the final test was administered. This test existed of two pages. The questions on the first page (10 retention questions and four comprehension questions) had to be completed and handed in before the pupils received the second page with the last comprehension question (the sorting question). This was done because the sorting question might give away answers to the previous questions. A maximum of 25 min was allowed for taking the test. When finished, pupils received a puzzle sheet so as not to disturb other pupils.

This procedure was repeated for all three texts. The total experiment lasted 6 weeks and included three instruction moments, nine practice sessions, and three final test moments (see Table 2).

Data-Scoring

Separate scoring forms were designed for the comprehension and factual retention questions. For each correct answered comprehension question, pupils could score 1 (partially correct answer) or 2 (completely correct answer), with a maximum of 10. For each correct answered retention question, pupils received a score of 1, with a maximum of 10. Spelling or grammatical errors were ignored.

To determine whether the quality of pupils' retrieval practice answers and pupils' summaries improved during practice sessions, data from the practice sessions were collected. The retention questions from the practice sessions were scored in the same way as the final test questions. For the data analysis, scores were converted to percentages. To score the quality of the summaries, the expert summary (described in the material section) was divided into idea units. The score of the quality of the pupil's summary was expressed as the percentage of these idea units that appeared in the summary.

The second author and an independent assessor scored 12.3% of the retention questions from the practice sessions; 12.3% of the summary idea units; 11.7% of the final retention questions, and 11.7% of the final comprehension questions. The interrater correlation between the two assessors was, respectively: .98, .96, 1,00, and .96, using Cohen's Kappa as an effect size (Cohen, 1960). Because of the high agreement, the second author scored the remaining tests.

The final retention questions and the final comprehension questions were controlled for content validity. The internal consistency of both the retention questions ($\alpha = .86$) and the comprehension questions ($\alpha = .76$) indicate a reliable measure.

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study, and follow Journal Artical Reporting Standards (Kazak, 2018). The ethical criteria of an institutional review board in the south of the Netherlands regarding information, recruitment, privacy, and data storage were met. All data and research materials were collected in 2016 and are available at the Data Archiving and Networked Services Easy repository, https://doi.org/10.17026/dans-2br-sj3y. Data were analyzed using IBM SPSS Statistics (Version 27, 2020). This study's design and its analysis were not preregistered.

Results

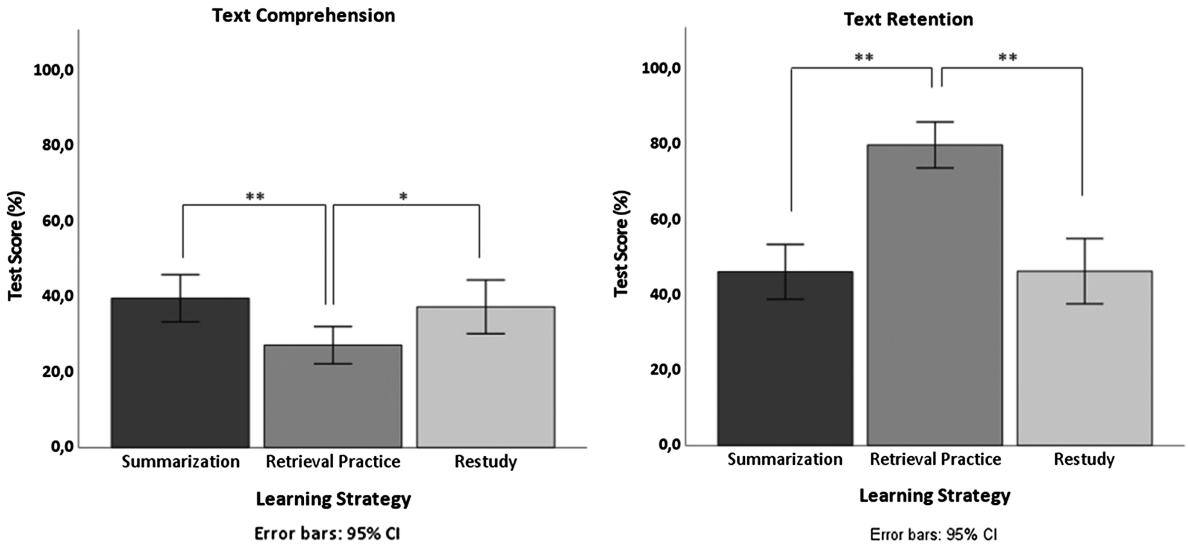
Results of the Final Tests

To analyze our research question about the effect of summarization, retrieval practice and restudy on text comprehension and factual text retention, two repeated measures analyses of variance (ANOVAs) were conducted. One with the independent variable "learning strategy" and the dependent variable "test score comprehension", and one with the independent variable "learning strategy" and the dependent variable "test score retention". The independent variable had three levels: summarization, retrieval practice, and restudy. For the repeated measures ANOVA on the comprehension questions, the assumption of sphericity was met. The repeated measures ANOVA on the text comprehension questions showed a statistically significant effect of learning strategy, $F(2, 112) = 7.77$, $p < .001$, $\eta_p^2 = .12$. This is a medium-large effect according to Cohen's threshold values for effect sizes (Cohen, 1988). A post hoc test revealed that, on the comprehension questions, there was no statistically significant difference between the learning strategies summarization ($M = 39.65$, $SD = 23.60$) and restudy ($M = 37.37$, $SD = 26.89$), $p = 1.00$. The effect of the learning strategy retrieval practice ($M = 27.19$, $SD = 18.67$) was significantly smaller than the effect of the learning strategies summarization, $p < .001$, and restudy, $p = .011$. Figure 1 represents these results.

For the repeated measures ANOVA on the retention questions, the assumption of sphericity was violated, $\chi^2(2, N = 57) = 8.21$, $p = .017$, $\varepsilon = .90$, therefore a Huyn-Feldt correction was applied (Huynh & Feldt, 1976). The repeated measures ANOVA on the text retention questions showed a statistically significant effect of learning strategy, $F(1.81, 101.30) = 42.31$, $p < .001$, $\eta_p^2 = .43$. This is a large effect according to Cohen's threshold values for effect sizes (Cohen, 1988). A post hoc test revealed that, on the retention questions, the effect of the learning strategy retrieval practice ($M = 79.83$, $SD = 23.03$) was

Figure 1

Final Test Scores on Comprehension and Retention Questions



Note. Significant differences (* $p < .05$. ** $p < .001$) between the learning strategies summarization, retrieval, and restudy on the comprehension questions (left) and the retention questions (right). Test scores are in percentages. CI = confidence interval.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 3
The Mean Scores During the Retrieval Practice and Summarization Practice Sessions, Expressed in Percentages

Learning strategy	Practice Session 1 (*SD*)	Practice Session 2 (*SD*)	Practice Session 3 (*SD*)	Final retention test (*SD*)
Summarization	28.07 (19.09)	38.69 (20.70)	49.05 (21.58)	
Retrieval practice	47.37 (29.49)	63.51 (26.76)	77.19 (25.20)	79.83 (23.03)

Note. Summarization: The quality of the summary is expressed by the mean percentage of included idea units. Retrieval practice: The mean percentage of correct answers on the practice sessions tests is presented (*SD* in parentheses).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

significantly larger than the effect of the learning strategies summarization ($M = 46.14$, $SD = 27.44$), $p < .001$, and restudy ($M = 46.32$, $SD = 32.71$), $p < .001$. There was no statistically significant difference between the learning strategies summarization and restudy, $p = 1.00$. Figure 1 represents these results.

Results of the Summary Quality During Practice Sessions

The percentage of correct idea units reproduced in a pupil's summary represented the quality of the summary. Table 3 shows the quality of the summaries expressed in percentages. To compare the quality of the summaries over time, a repeated measures ANOVA was conducted. Because the assumption of sphericity was violated, $\chi^2(2, N = 57) = 7.24$, $p = .027$, $\varepsilon = .92$, a Huyn-Feldt correction was applied (Huynh & Feldt, 1976). A significant effect was found for the quality of the summaries, $F(1.84, 102.76) = 32.79$, $p < .001$, $\eta_p^2 = .37$. This is a large effect according to Cohen's threshold values for effect sizes (Cohen, 1988). Table 3 shows the increase over time. A post hoc test revealed that the quality of the summaries in the third practice session ($M = 49.05$, $SD = 21.58$) was significantly higher than the quality of the summaries in the second ($M = 38.69$, $SD = 20.70$), $p < .001$, and the first practice session ($M = 28.07$, $SD = 19.09$), $p < .001$). The quality of the summaries in the second practice session was significantly higher than the quality of the summaries in the first practice session ($p < .001$). There was a 21% increase of summary quality between the first and the last practice session.

Additionally, the score on the third summary is a significant predictor of the score on the comprehension test. We tested the impact of an idea unit being included in the third summary on the final comprehension test score (raw score, with a maximum of 10). The predicted increase in comprehension is .50 per included idea unit in the third summary, $\beta = .50$; $t(55) = 5{,}40$; $p < .001$. The score on the third summary also explains a significant portion of the variance in the comprehension test score, $R^2 = .35$, $F(1.55) = 29{,}18$; $p < .001$.

Results of Retrieval Practice Scores

Table 3 presents the test scores of the retrieval practice sessions and the final retention test expressed in percentages. To compare the scores between practice sessions and the final retention test, a repeated measures ANOVA was conducted. Because the assumption of sphericity was violated, $\chi^2(5, N = 57) = 38.85$, $p < .001$, $\varepsilon = .71$, a Greenhouse–Geisser correction was applied (Greenhouse & Geisser, 1959). A significant difference was found between the scores in the retrieval practice sessions, $F(2.13, 119.25) = 57.43$, $p < .001$, $\eta_p^2 = .51$. This is a large effect according to Cohen's threshold values for effect sizes (Cohen, 1988). Table 3 shows the increase over time. A post hoc test revealed that the test scores in the third practice session ($M = 77.19$, $SD = 25.20$) were significantly higher than the test scores in the second ($M = 63.51$, $SD = 26.76$), $p < .001$, and the first practice session ($M = 47.37$, $SD = 29.49$), $p < .001$. The test scores of the second practice session were significantly higher than the test scores of the first practice session, $p < .001$. Finally, there was no statistically significant difference between the test scores on the third practice session and the final retention test ($M = 79.83$, $SD = 23.03\%$), $p = 1.00$. There was a 30% increase in test scores between the first and the last practice session.

Discussion

This research examined the effects of summarization and factual retrieval practice on text comprehension and text retention in an authentic elementary school setting, using a restudy control condition. Teachers executed the experiment and used authentic course materials. Below, the effects on text comprehension and text retention are discussed separately.

The Effects of Summarization and Factual Retrieval Practice on Text Comprehension

In our study, the effect of summarization on text comprehension was stronger compared to retrieval practice using factual knowledge questions. However, there was no benefit for summarization compared to the learning strategy restudy (control condition). The stronger effect of summarization and restudy on text comprehension compared to retrieval practice is in line with the theory that writing a summary or restudying a text increases a reader's focus on the main thoughts of that text, allowing the reader to establish causal, logical and explanatory relations between different parts of the text, which leads to better text comprehension (Kintsch, 1988; Marzano et al., 2001; Wade-Stein & Kintsch, 2004). This positive effect on text comprehension may not occur when using factual retrieval questions, because the focus of the reader shifts from the main thoughts and the interrelationships in the text, to the text's facts and details.

The limited effect of factual retrieval practice on text comprehension is in line with previous research with adolescents by Agarwal (2019) and McDaniel et al. (2013), who found no benefit of factual retrieval practice on higher order or application questions. However, in his first and second experiment, Butler (2010) did find a positive effect of factual retrieval practice questions on new inferential questions in undergraduate students. The contrast between our results and Butler's (2010) findings could be explained by a difference in the level of relatedness between our retention questions and our comprehension questions administered in the final test. In Butler's (2010) experiments, factual retention questions that were practiced with, were highly related to the final test inference questions, while in our study retention questions and comprehension questions were not highly related.

Hence, our findings do not necessarily imply that retrieval practice in general does not aid text comprehension. If, during practice sessions, retention questions are used that are highly related to final test inference questions or retention questions are used that focus on interrelationships in a text, these questions may very well enhance a pupil's text comprehension, as supported by research in adolescents (e.g., Agarwal, 2019; Butler, 2010).

The absence of a significant difference between the effect of the learning strategies summarization and restudy on text comprehension is noteworthy. The extra effort, that is necessary to produce a summary that describes the main thoughts of a text, did not lead to better text comprehension than simply restudying the text. An explanation for this finding might lie in the quality of the summaries that pupils at this age are able to produce. Although, the quality of the summaries, written after three summarization practice sessions, was positively correlated with the scores on the final test comprehension questions, on average, our participants only managed to include half of the idea units in their summaries. These findings show that pupils of this age struggle to write the high-quality summary that is needed to gain positive effects on text comprehension. This finding is in line with research of Brown et al. (1983), showing that younger pupils struggle to write a high-quality summary, compared to college students and high school students. Brown et al. (1983) argued that the possibility to write a high-quality summary is age-related, because it is not just a result of retention, but requires judgment, effort, knowledge, and strategies. Third- and fourth-grade elementary school pupils may lack the metacognitive skills to be aware of the structure and the relatedness of ideas within a text (Baker & Brown, 1984; Gogtay et al., 2004), and are therefore not able to write the high-quality summary needed to improve text comprehension. These arguments are supported by research of Brod (2021), in which it is argued that, compared to older students, elementary school pupils are less able to apply learning strategies effectively and efficiently due to a limited development of their metacognitive skills. These findings imply that training programs to increase a pupils' summary quality may be necessary for the learning strategy summarization to be beneficial compared to the learning strategy restudy.

The Effects of Summarization and Factual Retrieval Practice on Text Retention

In the current experiment, retrieval practice was a more effective learning strategy than summarization and restudy to retain factual knowledge from a text. These results indicate that pupils practicing with factual retrieval practice questions may have placed a stronger emphasis on the factual knowledge in the text, whereas pupils that wrote a summary or restudied the text focused on the main thoughts of that text instead of on the facts and details (Brown et al., 1983; Hagaman et al., 2016; Hidi & Anderson, 1986; Oded & Walters, 2001). These results are also in line with previous findings in other fields of elementary education, where retrieval practice led to better retention of factual knowledge compared to restudy (Goossens et al., 2014; Jones et al., 2016; Karpicke et al., 2016).

It could be argued that the advantage of retrieval practice compared to restudy was caused by the greater amount of time on task in the retrieval practice sessions (10 min of restudy vs. 25 min of retrieval practice). However, during the summarization practice sessions, pupils also restudied the text for 10 min, but additionally wrote a summary for 20 min, and finally received feedback for 10 min (40 min for summarization vs. 25 min for retrieval practice). Despite this disadvantage of 15 min in terms of time on task, practice with factual retrieval questions still had a larger effect on text retention. This shows that giving additional time did not lead to better performance. Also given the fact that adding reading sessions shows diminishing efficiency (Rothkopf, 1968), it is unlikely that the difference between the effects of retrieval practice and restudy was caused by the difference in time on task.

Future Research and Limitations

We suggested that the relatively low impact of writing a summary compared to restudy on text comprehension could be explained by the poor quality of pupils' summaries. It is possible that summarization may be more effective than restudy when this quality is enhanced. Therefore, it would be interesting to investigate if training programs, such as the structure strategy training of Meyer (see Meyer & Ray, 2011, for a review), can be used to improve the quality of elementary school pupils' summaries and thereby their text comprehension in an authentic classroom setting. Also, it could be interesting to investigate if certain mistakes in summaries are predictive of the score on comprehension questions. This could aid method developers in obtaining a deeper understanding of the areas that necessitate further practice to enhance the summarization skills of elementary school pupils.

Additionally, the retention questions in our study were factual knowledge questions. In future research, it could be examined if other forms of retrieval practice questions could enhance the text comprehension of elementary school pupils, following up on previous research of Agarwal (2019), Butler (2010), Karpicke and Blunt (2011), and McDaniel et al. (2013).

Finally, it is noteworthy that during the process of learning, often multiple goals have to be achieved. Therefore, in order to create an optimal learning environment it might be most effective to combine different learning strategies (see Roelle et al., 2022). In our study, when learning from text materials, the goal of text comprehension is enhanced by the learning strategies summarization and restudy, whereas the goal of text retention is enhanced by the learning strategy retrieval practice. Because of the importance of both goals, it might be most effective when different learning strategies are combined within one practice session. Hence, future research should investigate combinations of learning strategies to optimize learning.

Constraints on Generality

The results in our study are generalizable to 8 to 10-year-old pupils using texts with the same degree of difficulty as the texts we used in this research. Results may not be generalizable to other age groups because of a difference in reading stage and metacognitive skills.

Conclusion

This study demonstrated that in an authentic elementary school setting, even without extensive training, writing a summary had a larger positive effect on third- and fourth-grade pupils' text comprehension compared to retrieving factual knowledge from a text. However, writing a summary did not have a larger effect on text comprehension than simply restudying that text. Thus, it is

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

important for teachers to understand that writing a summary does not automatically improve text comprehension compared to the more time-efficient learning strategy restudy. To be effective, a summary may need to be of high quality, which appears to be difficult to achieve for pupils of this age. When it comes to factual text retention, answering factual questions about a text after initial reading proves more effective for pupils in retaining this knowledge for the long term than either writing a summary of the text or rereading it.

In conclusion, this study demonstrates that a learning session often includes multiple goals. Based on our findings, it can be argued that employing diverse learning strategies is of importance for effectively achieving each individual goal.

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, *87*(3), 659–701. https://doi.org/10.3102/0034654316689306

Agarwal, P. K. (2019). Retrieval practice & Bloom's taxonomy: Do students need fact knowledge before higher order learning? *Journal of Educational Psychology*, *111*(2), 189–209. https://doi.org/10.1037/edu0000282

Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: A systematic review of applied research in schools and classrooms. *Educational Psychology Review*, *33*(4), 1409–1453. https://doi.org/10.1007/s10648-021-09595-9

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 353–394). Longman. Retrieved September 14, 2023, from https://shorturl.at/enF69

Brod, G. (2021). Generative learning: Which strategies for what age? *Educational Psychology Review*, *33*(4), 1295–1318. https://doi.org/10.1007/s10648-020-09571-9

Brown, A. L., Campione, J. C., & Day, J. D. (1981). Learning to learn: On training students to learn from texts. *Educational Researcher*, *10*(2), 14–21. https://doi.org/10.2307/1174401

Brown, A. L., Day, J. D., & Jones, R. S. (1983). The development of plans for summarizing texts. *Child Development*, *54*(4), 968–979. https://doi.org/10.2307/1129901

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1118–1133. https://doi.org/10.1037/a0019902

CED-groep, Centrum Educatieve Dienstverlening. (2022). *Nieuwsbegrip-over Nieuwsbegrip*. Retrieved December 31, 2023, from https://www.nieuwsbegrip.nl/over-nieuwsbegrip

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Chall, J. (1983). *Stages of reading development*. Retrieved July 5, 2022, from http://www.landmarkoutreach.org/wp-content/uploads/Challs-Stages-of-Reading-Development.pdf

Chan, J. C., McDermott, K. B., & Roediger, H. L., III. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General*, *135*(4), 553–571. https://doi.org/10.1037/0096-3445.135.4.553

Christie, F. (2010). The ontogenesis of writing in childhood and adolescence. *The Routledge international handbook of English, language and literacy teaching* (pp. 146–158). Routledge. Retrieved September 14, 2023, from https://shorturl.at/hjktv

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Erlbaum. https://doi.org/10.1016/0198-9715(90)90050-4

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., III, Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L., & Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(21), 8174–8179. https://doi.org/10.1073/pnas.0402680101

Goossens, N. A., Camp, G., Verkoeijen, P. P., & Tabbers, H. K. (2014). The effect of retrieval practice in elementary school vocabulary learning. *Applied Cognitive Psychology*, *28*(1), 135–142. https://doi.org/10.1002/acp.2956

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. https://doi.org/10.1007/BF02289823

Hagaman, J. L., Casey, K. J., & Reid, R. (2016). Paraphrasing strategy instruction for struggling readers. *Preventing School Failure*, *60*(1), 43–52. https://doi.org/10.1080/1045988X.2014.966802

Hidi, S., & Anderson, V. (1986). Producing written summaries: Task demands, cognitive operations, and implications for instruction. *Review of Educational Research*, *56*(4), 473–493. https://doi.org/10.3102/00346543056004473

Huynh, H., & Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, *1*(1), 69–82. https://doi.org/10.3102/10769986001001069

International Business Machines Corp. (2020). *IBM SPSS statistics for windows* (Version 27.0).

Jones, A. C., Wardlow, L., Pan, S. C., Zepeda, C., Heyman, G. D., Dunlosky, J., & Rickard, T. C. (2016). Beyond the rainbow: Retrieval practice leads to better spelling than does rainbow writing. *Educational Psychology Review*, *28*(2), 385–400. https://doi.org/10.1007/s10648-015-9330-6

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*(6018), 772–775. https://doi.org/10.1126/science.1199327

Karpicke, J. D., Blunt, J. R., & Smith, M. A. (2016). Retrieval-based learning: Positive effects of retrieval practice in elementary school children. *Frontiers in Psychology*, *7*, Article 350. https://doi.org/10.3389/fpsyg.2016.00350

Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist*, *73*(1), 1–2. https://doi.org/10.1037/amp0000263

Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, *29*(1), 10–16. https://doi.org/10.1111/ldrp.12025

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*(4), 294–303. https://doi.org/10.1037/0003-066X.49.4.294

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 989–998. https://doi.org/10.1037/a0015729

Lipko-Speed, A., Dunlosky, J., & Rawson, K. A. (2014). Does testing with feedback help grade-school children learn key concepts in science? *Journal of Applied Research in Memory and Cognition*, *3*(3), 171–176. https://doi.org/10.1016/j.jarmac.2014.04.002

Marsh, E. J., Lozito, J. P., Umanath, S., Bjork, E. L., & Bjork, R. A. (2012). Using verification feedback to correct errors made on a multiple-choice test. *Memory*, *20*(6), 645–653. https://doi.org/10.1080/09658211.2012.684882

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Marzano, R. J., Pickering, D., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Association for Supervision and Curriculum Development.

McDaniel, M. A., Howard, D. C., & Einstein, G. O. (2009). The read-recite-review study strategy: Effective and portable. *Psychological Science*, *20*(4), 516–522. https://doi.org/10.1111/j.1467-9280.2009.02325.x

McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., & Roediger, H. L. (2013). Quizzing in middle-school science: Successful transfer performance on classroom exams. *Applied Cognitive Psychology*, *27*(3), 360–372. https://doi.org/10.1002/acp.2914

McDermott, K. B., Agarwal, P. K., D'Antonio, L., Roediger, H. L., III, & McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *Journal of Experimental Psychology: Applied*, *20*(1), 3–21. https://doi.org/10.1037/xap0000004

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*(1), 1–43. https://doi.org/10.1207/s1532690xci1401_1

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation*, *51*, 297–384. https://doi.org/10.1016/S0079-7421(09)51009-2

Meyer, B. J., & Ray, M. N. (2011). Structure strategy interventions: Increasing reading comprehension of expository text. *International Electronic Journal of Elementary Education*, *4*(1), 127–152. Retrieved September 14, 2023, from https://www.iejee.com/index.php/IEJEE/article/view/217

Mok, W. S. Y., & Chan, W. W. L. (2016). How do tests and summary writing tasks enhance long-term retention of students with different levels of test anxiety? *Instructional Science*, *44*(6), 567–581. https://doi.org/10.1007/s11251-016-9393-x

Mulder, L., & Meijnen, W. (2013). *Onderwijsachterstanden in de BOPO-periode, 2009–2012*. Een review, 2009-2011. Retrieved November 27, 2023, from https://www.nro.nl/sites/nro/files/migrate/Reviewstudie-Onderwijsachterstanden.pdf

National Reading Panel, National Institute of Child Health, Human Development, National Reading Excellence Initiative, National Institute for Literacy, & United States Department of Health. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health. Retrieved June 29, 2022, from https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

New South Wales Centre for Effective Reading. (2022). *Comprehension-Integrated strategies*. Retrieved June 29, 2022, from https://cer.schools.nsw.gov.au/content/dam/doe/sws/schools/c/cer/localcontent/integrated_strategies.pdf

Nguyen, K., & McDaniel, M. A. (2016). The JOIs of text comprehension: Supplementing retrieval practice to enhance inference performance. *Journal of Experimental Psychology: Applied*, *22*(1), 59–71. https://doi.org/10.1037/xap0000066

Nurhayati, D. A. W., & Fitriana, M. W. (2018). Effectiveness of summarizing in teaching reading comprehension for EFL students. *Indonesian Journal of Language Teaching and Linguistics*, *3*(1), 33–50. https://doi.org/10.30957/ijoltl.v3i1.403

Oded, B., & Walters, J. (2001). Deeper processing for better EFL reading comprehension. *System*, *29*(3), 357–370. https://doi.org/10.1016/S0346-251X(01)00023-9

Organization for Economic Cooperation and Development. (1999). *Measuring student knowledge and skills: A new framework for assessment*. Retrieved December 31, 2021, from https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33693997.pdf

Pakzadian, M., & Rasekh, A. E. (2012). The effects of using summarizing strategies on Iranian EFL learners' reading comprehension. *English Linguistics Research*, *1*(1), 118–125. https://doi.org/10.5430/elr.v1n1p118

Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 3–8. https://doi.org/10.1037/0278-7393.31.1.3

Rinehart, S. D., Stahl, S. A., & Erickson, L. G. (1986). Some effects of summarizing training on reading and studying. *Reading Research Quarterly*, *21*(4), 422–438. https://doi.org/10.2307/747614

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roelle, J., Schweppe, J., Endres, T., Lachner, A., von Aufschnaiter, C., Renkl, A., Eitel, A., Leutner, D., Rummer, R., Scheiter, K., & Vorholzer, A. (2022). Combining retrieval practice and generative learning in educational contexts. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *54*(4), 142–150. https://doi.org/10.1026/0049-8637/a000261

Rothkopf, E. Z. (1968). Textual constraint as function of repeated inspection. *Journal of Educational Psychology*, *59*(1), 20–25. https://doi.org/10.1037/h0025378

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Rowley, T., & McCrudden, M. T. (2020). Retrieval practice and retention of course content in a middle school science classroom. *Applied Cognitive Psychology*, *34*(6), 1510–1515. https://doi.org/10.1002/acp.3710

Spirgel, A. S., & Delaney, P. F. (2016). Does writing summaries improve memory for text? *Educational Psychology Review*, *28*(1), 171–196. https://doi.org/10.1007/s10648-014-9290-2

Stevens, E. A., Park, S., & Vaughn, S. (2018). Summarizing and main idea interventions for upper elementary and secondary students: A meta-analysis of research from 1978 to 2016. *Remedial and Special Education*, *40*(3), 131–149. https://doi.org/10.1177/0741932517749940

Taylor, B. M. (1982). Text structure and children's comprehension and memory for expository material. *Journal of Educational Psychology*, *74*(3), 323–340. https://doi.org/10.1037/0022-0663.74.3.323

Thiede, K. W., Redford, J. S., Wiley, J., & Griffin, T. D. (2012). Elementary school experience with comprehension testing may influence metacomprehension accuracy among seventh and eighth graders. *Journal of Educational Psychology*, *104*(3), 554–564. https://doi.org/10.1037/a0028660

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press. Retrieved September 14, 2023, from https://ap.lc/oduIv

Wade-Stein, D., & Kintsch, E. (2004). Summary street: Interactive computer support for writing. *Cognition and Instruction*, *22*(3), 333–362. https://doi.org/10.1207/s1532690xci2203_3

Westby, C., Culatta, B., Lawrence, B., & Hall-Kenyon, K. (2010). Summarization expository texts. *Topics in Language Disorders*, *30*(4), 275–287. https://doi.org/10.1097/TLD.0b013e3181ff5a88

Wormeli, R. (2004). *Summarization in any subject: 50 techniques to improve student learning*. Association for Supervision and Curriculum Development. Retrieved September 14, 2023, from https://shorturl.at/yQRVW

Received April 3, 2023
Revision received September 25, 2023
Accepted October 24, 2023 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.