



## Retrieval-based learning versus discussion; which review practice will better enhance primary school students' knowledge of scientific content?

Anna Jakobsson, Jenny Loberg & Maria Kjörk

**To cite this article:** Anna Jakobsson, Jenny Loberg & Maria Kjörk (2024) Retrieval-based learning versus discussion; which review practice will better enhance primary school students' knowledge of scientific content?, *International Journal of Science Education*, 46:12, 1216-1238, DOI: [10.1080/09500693.2023.2283906](https://doi.org/10.1080/09500693.2023.2283906)

**To link to this article:** <https://doi.org/10.1080/09500693.2023.2283906>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 24 Nov 2023.



[Submit your article to this journal](#)



Article views: 1738



[View related articles](#)



[View Crossmark data](#)

# Retrieval-based learning versus discussion; which review practice will better enhance primary school students' knowledge of scientific content?

Anna Jakobsson <sup>a</sup>, Jenny Loberg <sup>b,c</sup> and Maria Kjörk <sup>a</sup>

<sup>a</sup>Department of Social and Behavioural Studies, Division of Educational Science and Languages, University West, Trollhättan, Sweden; <sup>b</sup>Foundation Nordens Ark, Åby Säteri, Hunnebostrand, Sweden; <sup>c</sup>Department of Animal Environment and Health, Swedish University of Agricultural Science, Skara, Sweden

## ABSTRACT

Retrieval-based learning, using tests for content review, frequently proves more effective for knowledge retention compared to alternative methods. Extensive research has explored this with older students, often in contrast to more passive techniques like rereading or note rewriting, typically focusing on vocabulary content, in non-classroom settings and assessing knowledge retention within a week. In this study we conducted a classroom experiment to compare the efficacy of retrieval-based learning to another active method of reviewing content, namely discussion, in promoting long-term knowledge retention among primary school science students. Additionally, we assessed the students' perceptions of these reviewing methods. A total of eighty-one primary school students participated in the study, which encompassed a lesson conducted at a zoo, followed by a review of the lesson content using either retrieval-based learning or discussion. Our findings indicated no impact of the method of reviewing on knowledge retention after either one week or four weeks. Nevertheless, notable effects were observed in terms of motivation, as students who employed retrieval-based learning for reviewing reported heightened levels of interest and enjoyment, along with perceiving greater benefits from the process. Additionally, boys in the discussion group reported increased curiosity, in contrast to girls who reported an inverse pattern.

## ARTICLE HISTORY

Received 24 March 2023  
Accepted 12 November 2023

## KEYWORDS

Retrieval-based learning;  
primary school; motivation

## Introduction

Learning experiences in primary science that have been demonstrated to enhance students' knowledge encompass a diverse array of activities and methods. For instance, outdoor teaching (as demonstrated by Wünschmann et al., 2017), inquiry-based learning (as exemplified by Akaygun & Adadan, 2021), and taking part in science discourse by discussion and talking (as reviewed by Bae et al., 2021) are all noteworthy examples. In this study,

**CONTACT** Anna Jakobsson  [anna.jakobsson@hv.se](mailto:anna.jakobsson@hv.se) 

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

we focus on a method known as retrieval-based learning for the purpose of reviewing scientific content subsequent to an initial learning encounter. This practice stems from the field of cognitive research where an array of methodologies has exhibited strong favourable effects on the process of learning and memory, as for example altering between different study content (interleaving), study information at strategically placed intervals (spacing), and active recall of content from memory (retrieval-based learning) (e.g. Bjork, 1975, 2011; Dunlosky et al., 2013). These methods have been claimed to be generally underutilised in teaching (Didau, 2015; Dunlosky et al., 2013). However, cognitive aspects of learning, as for example interleaving, spacing and retrieval based-learning, are increasingly being incorporated into for example math education (Barton, 2018), medical education (Schmidt & Mamede, 2020) and primary science (Earle & McMahon, 2022). Additionally, in the last decade, educational research and cognitive research have reunited (Agarwal et al., 2012), as demonstrated by studies such as those conducted by Wiklund-Hörnqvist (2014), which combine classroom studies and laboratory experiments.

### ***Retrieval-based learning and the test-effect***

In the present study, our focus is on retrieval-based learning (Karpicke et al., 2016), also referred to as test-enhanced learning (Jaeger et al., 2015). The core principle of this method lies in students actively recalling content, whereby the cognitive effort expended enhances their ability to remember the content in subsequent instances (Karpicke & Grimaldi, 2012). It's important to note that this recall is intended to facilitate learning rather than serve as an assessment of the students' knowledge. The form of active recall, whether it entails answering written questions or responding to oral questions posed by a teacher, is not significant. A common scenario in the science classroom could entail the teacher, towards the conclusion of the lesson, prompting students to try and recall the content on their own, without consulting their notes or participating in discussions with their peers.

The so called 'test-effect' refers to the positive impact on knowledge retention that results from retrieval-based learning compared to other methods of studying, and is well documented in cognitive psychology (Bjork, 1975; Roediger & Karpicke, 2006). A typical experimental setup involves students reading a text or a word list and then re-studying the information either by further reading or by active recall in a test. Even with equal amounts of time spent on the content, students reviewing through testing typically perform better on the final exam compared to those who review by further reading (as reviewed by Roediger & Butler, 2011). In other words, knowledge retention seems to be more favourable with a 'study-test-test-final exam' approach than with a 'study-study-study-final exam' approach.

What mechanism can explain the test-effect? A commonly used model characterises information stored in memory based on two factors: storage strength and retrieval strength (Bjork, 2011). Storage strength reflects how well information is associated with other memories and skills, while retrieval strength reflects the accessibility of the information, which depends on the recency of usage or availability of clues that remind the individual of the information. When memories are used, both storage strength and retrieval strength change, often they increase (Bjork, 2011). One plausible explanation for the test-effect is that the cognitive effort involved in trying to retrieve a memory makes it easier to access the next time (i.e. facilitates retrieval) and also

enhances its interconnectedness with other memories (i.e. increases storage strength) (Karpicke & Grimaldi, 2012). This is supported by studies showing that greater cognitive effort leads to a larger test-effect. For example, repetition through multiple-choice questions has been shown to elicit less or no test-effect compared to free-response questions (Kang et al., 2007).

### ***The relevance of retrieval-based learning for educational practice***

#### ***Meaningful learning and learning in a long-term perspective***

Retrieval-based learning has been criticised for promoting rote learning, to simply mirror the similarity between the test situation and the study situation and to be unable to measure meaningful learning (e.g. Karpicke & Blunt, 2011; Rohrer et al., 2010). While it is true that most studies of retrieval-based learning utilise final exam questions that are identical to the study questions, there are exceptions. For example, retrieval-based learning resulted in higher performance on the final exam compared to re-studying through the construction of a concept map, even when the final exam required concept mapping (Karpicke & Blunt, 2011). Also, the capacity for knowledge transfer, the ability to apply knowledge acquired in one context in a different context (e.g. Perkins & Salomon, 1992), has been demonstrated to be enhanced by retrieval-based learning (Rohrer et al., 2010).

Most studies on retrieval-based learning assess knowledge within a week, which can be considered relatively short when considering a long-term perspective of learning. The significance of time when evaluating different methods was demonstrated in a study involving medical students who were learning neuroanatomy (Sanders et al., 2019). When knowledge was assessed after one week the group that had been exposed passively to the content, with the lecturer showing and naming anatomical structures, showed better performance compared to the retrieval-based learning group. However, a month after the intervention, the opposite pattern was observed. A similar but shorter-term pattern was shown for psychology students where repeated reading resulted in better performance when the final test was immediately administered, whereas retrieval-based learning led to better performance when the test was given a week after the last content review (Lim et al., 2015).

Hence, multiple studies suggest that retrieval-based learning may enhance meaningful learning and learning in a long-term perspective. However, additional research would benefit from exploring impacts on knowledge transfer and incorporating assessments of knowledge at various time points following the intervention, including durations exceeding a week.

#### ***Classroom studies and active compared to passive methods***

Although the majority of studies investigating retrieval-based learning have been conducted in laboratory settings (Roediger & Butler, 2011), there is now a growing body of research conducted in classrooms (for review see Moreira et al., 2019; Ortega-Tudela et al., 2021). This is essential to assess the method's relevance for educational practice. However, in the literature review of retrieval-based learning conducted in classrooms by Moreira et al. (2019), most studies used passive reviewing of content (as for example re-reading or re-writing notes) or no reviewing at all, as compared to the retrieval-

based method. From a pedagogical standpoint, it is vital to compare retrieval-based learning to other active methods. Education in general, and scientific teaching in particular, has a long-standing tradition of designing review situations in which students are active, and the significance of generative learning activities such as summarising, explaining or taking part in discussion is well established (Ausubel, 1968; Bae et al., 2021; Hattie, 2009;). It is still unclear how other active methods differ from retrieval-based learning. Nonetheless, some studies have demonstrated that retrieval-based learning is more effective than other active methods, such as active generation of word pairs when learning vocabulary (Karpicke & Zaromb, 2010), and the aforementioned study by Karpicke and Blunt (2011), which revealed the superiority of retrieval-based learning over the creation of a concept map.

### *Younger students and science learning in classroom settings*

Although most studies on retrieval-based learning have concerned adults and older students (Roediger & Butler, 2011; Roediger & Karpicke, 2006), positive effects have also been shown for preschool children and for primary and middle school students (review by Fazio & Marsh, 2019; Fritz et al., 2007; Jaeger et al., 2015; Karpicke et al., 2016). Most of these studies, however, are made in laboratory settings. In a review by Moreira et al. (2019) of retrieval-based learning in classroom settings, only five studies involved students between the ages of 7 and 13 years. Among these, three studies were centred on the acquisition of scientific content (Lipko-Speed et al., 2014; McDaniel et al., 2011; McDaniel et al., 2013).

In their study on retrieval-based learning through quizzes with 8th grade students, McDaniel et al. (2011) found a positive effect of multiple quizzes on the retention of content from genetics, evolution, and anatomy. However, the quizzes were not compared to another method of reviewing content, but students were given tests on half of the content that was later assessed in a final test. The retention of content included in the quizzes were significantly higher than the content not included, and the benefits of quizzing persisted on cumulative semester and end-of-year exams. In a similar study on 9th grade students where assessment of knowledge took place after 24 h McDaniel et al. (2013) found a positive effect on retention of quizzing on content of cells, machines/energy, and animals. In the study by Lipko-Speed et al. (2014), on the other hand, retrieval-based reviewing of science content about light and sound in 5th graders was compared to another method of reviewing, namely re-reading. In their study 15 scientific key terms were practiced twice with either 'retrieval only', 'retrieval-plus-feedback' or 're-reading'. Knowledge was assessed in a final test two days after the last reviewing of terms. The results showed that the key terms reviewed by 'retrieval-plus-feedback' were better remembered than terms reviewed by 'retrieval only' or by 're-reading'. Rowley and McCrudden (2020) found that middle school students during learning of a text with scientific content benefited more from trying to remember as much as possible of the text than from copying their notes about the text. Knowledge was assessed in a test two days after the last reviewing of the text.

These results suggest that retrieval-based learning has the potential to be a complementary tool in primary and middle school science teaching. However, studies that compare the method of retrieval to other commonly used active methods as well as long-term learning in younger students are needed.

### ***Student experience and motivation***

An overall insight from cognitive research is that methods which are experienced as rewarding in teaching and learning situations are not always the ones best promoting long-term learning (e.g. Bjork & Bjork, 2014; Karpicke & Blunt, 2011; Karpicke & Grimaldi, 2012; Roediger & Karpicke, 2006). Rather, it seems as if circumstances effective for a short-term performance are of disadvantage for learning in the longer term, and those favouring long-term learning are of disadvantage for performance in the actual learning occasion. Paradoxically, this means that what intuitively feels good for both teachers and students, methods where the student learns quickly and the teaching runs smoothly, is not always what optimises learning in the longer term (Bjork & Bjork, 2014).

The discrepancy between students' subjective perception of a method's effectiveness and their actual learning has been repeatedly demonstrated in retrieval-based learning. For example, in a study by Karpicke and Grimaldi (2012), students who re-studied material through questions and active recall were found to underestimate their learning, whereas those who re-studied by repeatedly reading overestimated their learning. A possible explanation for this discrepancy could be that students form a false connection between the ease of processing content and the degree to which they have learnt it (Karpicke et al., 2009). When a student reads a text multiple times, the content becomes familiar and recognisable, creating a feeling of knowledge possession. In contrast, re-studying content by actively recalling and answering questions, often requires greater cognitive effort, creating a sense of resistance. Given that commitment and motivation are important prerequisites for students' learning (Bybee et al., 2006; Deci & Ryan, 1985), it is crucial to consider students' experience with the method when it comes to implementing it in the school. We know of no studies that have investigated this for retrieval-based learning, except for experiences mentioned above about self-estimation of learning. The experience of students can be of special importance in science education, where encouraging students' interest in the subject has a long tradition, and interest has been shown to affect learning outcomes (e.g. Toli & Kallery, 2021). Furthermore, research has indicated that the manifestation of talents and strengths within the school environment is influenced by gender (Määttä & Uusiautti, 2020), and interactions in scientific classrooms have been shown to vary between boys and girls (e.g. Eliasson et al., 2016; Jones & Wheatley, 1990). As a result, the experience of learning methods could potentially be influenced by gender, an aspect that, to the best of our understanding, has not been explored in the context of retrieval-based learning.

### ***Aim and research questions***

The aim of this study was to compare two active methods of reviewing scientific content, retrieval-based and discussion, among primary school students within a genuine teaching sequence. Discussion was selected as the active method for comparison with retrieval-based learning because it was a common practice for content review in the classes. The study's primary objective was to assess the effectiveness of retrieval-based learning in promoting knowledge retention in the complex setting of school teaching, over a period of four weeks. An additional objective was to evaluate the students' perceptions and experiences with both methods. While a method may prove effective in acquiring new

knowledge, it could have adverse consequences if it causes discomfort or unease among the students in the learning environment. Our aim was to address the following questions:

- (1) Is there a difference in effectiveness between retrieval-based learning and discussion as a review method for promoting long-term knowledge among primary school students, assessed at one week and four weeks after the review?
- (2) Do simpler forms of knowledge, such as facts and concepts, experience different impacts compared to more complex forms of knowledge, like contextual understanding, when subjected to these two review methods?
- (3) Does retrieval-based learning differ from discussion in its influence on students' ability of knowledge transfer?
- (4) Do primary school students' experience with the two review methods (retrieval-based learning vs discussion) differ in terms of interest, perceived benefit, pressure/effort, and perceived competence, and is this contrast influenced by gender?

## **Method**

### ***Participants***

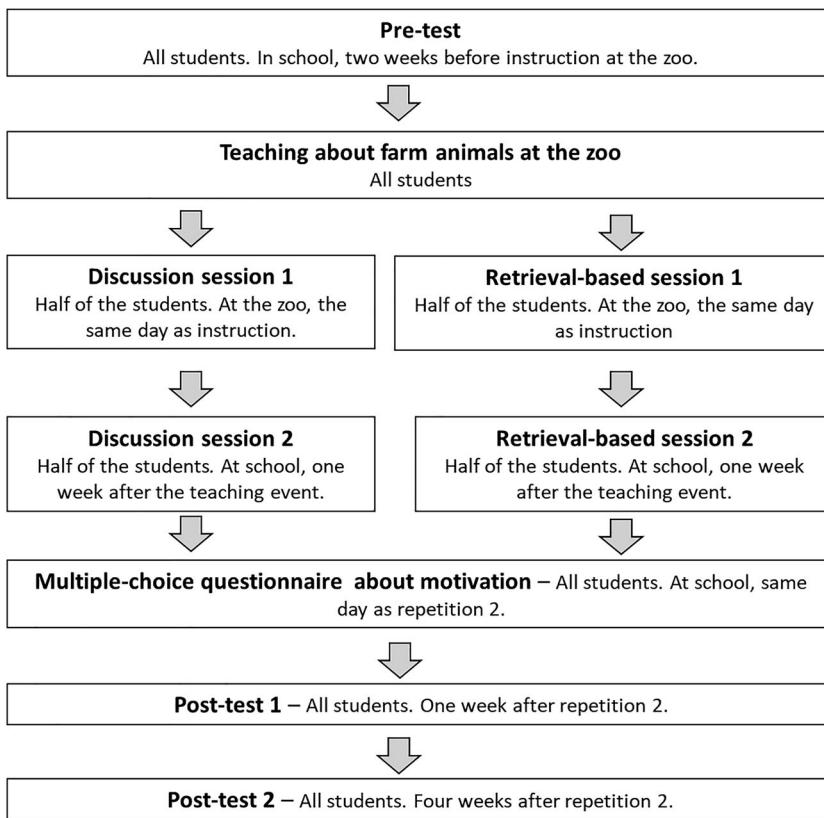
A total of 81 primary school students, aged 10–11 (48% girls), from two different schools participated in the study. Both schools were located in small towns in the countryside in southwest Sweden, and were chosen based on their proximity to Nordens Ark, a zoo where part of the study was conducted. We obtained consent from both the students and their parents, and provided information about the study's purpose, as well as the option for participants to withdraw from the study at any point. We also visited the classes after the study was finished, to inform students and teachers about results and the research process in general. The study followed ethical guidelines established by the Swedish Research Council and was approved for implementation by the Swedish Ethical Review Authority. Regarding disclosure, we have no competing interests to declare.

### ***Experimental design***

We used a randomised experimental design where primary school students were instructed on the topics of domestication and animal welfare in three different species of farm animals. The students then reviewed the content using either retrieval-based learning or discussion (Figure 1). Before the experiment, we administered a pre-test to assess the students' baseline knowledge. The same test was then given one week and one month after the final repetition. Additionally, the students answered a multiple-choice questionnaire about motivation.

### ***Teaching at about farm animals at the zoo***

The initial teaching instruction was conducted at Nordens Ark, a private non-profit foundation engaged in conservation, research and public education. Nordens Ark runs a farm that is specialised in maintaining native Swedish breeds of farm animals



**Figure 1.** The experimental design of the study.

and works on conservation breeding programmes. The students were taught about the domestication of pigs, chickens, and horses, as well as the animal welfare in these species. This content relates to the central content of evolution and natural farming in the Swedish science syllabus for grade 4–6 (Skolverket, 2022). The teaching was conducted by the authors and was carried out in small groups of students with accompanying teachers. Students had been randomly assigned to the teaching groups with regards to gender, with the aim of achieving gender-balanced groups. Initially, all students, both boys and girls, were randomly assigned into groups. Subsequently, a few randomly selected students were moved between groups if the gender distribution in the groups was imbalanced. The rationale behind this procedure was to prevent gender-related effects on the teaching situation (for a review of such aspects see Määttä & Uusiutti, 2020).

The instruction involved observing the living animals, listening to stories about domestication, viewing a picture board displaying various breeds of the animal, and talking about animal welfare. By rotating the students and keeping the person teaching about the animal consistent, the teaching was standardised so that all students received the same content and spent an equal amount of time with each animal. A total of four days of teaching was conducted at the zoo.



## Reviewing of content

Students were randomly assigned to review the content either through retrieval-based learning or discussion, with respect taken to gender and teaching group to avoid possible systematic effects of these factors, using the same procedure as describe above for the assignment to the teaching groups. The first session of reviewing content was conducted at the zoo after a lunch break, and the second session was held one week later at school (Figure 1). To mitigate any potential teacher effect, the persons who conducted the retrieval-based learning and discussion sessions were altered between the two occasions.

Both methods of reviewing content, retrieval-based and discussion, utilised the same questions with identical wording to reinforce the content. During the retrieval-based session, students were instructed to silently reflect on the questions and record their thoughts on paper, which was then collected at the end of the session. Informing students that their responses would be collected is important, as it can increase their motivation to make the cognitive effort required for memory recall (Bjork & Bjork, 2014). Moreover, to reduce possible stress and anxiety, we carefully informed the students that this was not a test to assess their knowledge. For each question, the students were encouraged to try to remember the specific instance when the content was covered. For some questions, a clue was given after a period of time, such as *'the word rhymes with ...'*. After each question, the correct answer was provided and there was a brief opportunity for students to ask questions. We deviated from the recommendation by Karpicke et al. (2016), to investigate the testing effect without feedback. This decision was made because the focus of our study was not solely on the testing effect, but rather on how it is applied in real school practice. In primary school, it is common to provide students with answers and feedback when working with questions in the classroom.

During the discussion session some questions were discussed in the whole group while others were discussed in smaller groups. The person in charge of the session monitored the discussion to ensure a close connection to the question and to ensure that proper facts and relationships were finally reached upon. This method of reviewing the content was chosen in cooperation with the teachers at the schools, as it was considered a common way of reinforcing subject content in science teaching in the class.

## Knowledge tests

To ensure that the questions were easily understood by students of this age and that there was not too much background knowledge, a pilot test was conducted on two classes in a school not further involved in the study. After the pilot, the test was refined by omitting some questions and reformulating others.

The pre-test to assess background knowledge was administered approximately two weeks prior to the teaching at the zoo. To make the pre-test more engaging for the students, some additional and easy questions were added to the pre-test (for example, 'what is the name of the pigs' tail?') but were not included in the subsequent post-tests. Otherwise, the pre-test and the two tests after the teaching contained identical questions. The test included both multiple choice questions and open-ended questions (Appendix 1). Both questions as well as answer alternatives in multiple-choice questions were presented in different orders in the pre-test and the two post-tests. The test questions concerned

same content as the questions used during the reviewing sessions, but they were rephrased to avoid an exact similarity between the reviewing sessions and the test. A question about rabbit breeds not covered in the reviewing sessions was added to the tests to assess the student's ability of knowledge transfer. The first knowledge assessment test was conducted one week after the final reviewing of content, and the second test was conducted approximately one month later (Figure 1). Students took the tests in mixed groups, with respect to their method of reviewing content, to avoid possible classroom effects.

During the evaluation of the tests, a standard evaluation template was constructed for the open-ended questions, and tests from the first school were evaluated by two of the authors to ensure the reliability of the template. For tests' from the second school, the evaluation template was used, and the two individuals consulted with each other in case of ambiguous answers. Furthermore, questions were classified as either simple or complex, following the same method by Karpicke and Blunt (2011). Simple questions could be answered without comprehending the interrelationship with other knowledge, whereas complex questions entailed multiple components that had to be linked. Some questions were excluded from the analysis of simple versus complex content, due to difficulties in classifying into either category (see Appendix 1 for classification).

### **Motivation**

The students' experience of the methods of reviewing content was investigated through a multiple-choice questionnaire answered directly after the last session at school. The questions followed a Likert scale and concerned interest, stress, experienced benefit, and experienced competence (Appendix 2). The questionnaire was designed by adapting a short scale for intrinsic motivation used for evaluation of students' experience of out-of-school teaching by Wilde et al. (2009). Questions 1–9 were translated from 'Short scale for intrinsic motivation Inventory' by Wilde et al. (2009), and questions 10–12 were constructed with reference to Mellor and Moore (2014) and Taherdoost (2019). We also added a question about whether the student wanted to use the method of reviewing content more often in school (question 13) and a question about curiosity and acquiring new questions (question 14). One weakness of the Likert scale is the risk of respondents adjusting their answers to please the researcher (Taherdoost, 2019). To mitigate this risk, we informed the students that we were genuinely interested in their experiences and that we would not be offended if they found the method boring or unsuitable for learning.

### **Statistical analysis**

Statistical analyses were conducted using SPSS 27. General linear models were used to analyse the effects on knowledge and motivation, with gender and method of reviewing as fixed factors, school as a random factor, and interaction effects for gender\*method of reviewing, and school\*treatment. Interaction terms with  $p > 0.25$  were excluded from the final models (Underwood, 1997). The homogeneity of variances was assessed using Levene's test. Analyses of knowledge were made on (i) total score (ii) scores on simple content (iii) scores on complex content, and (iv) ability of knowledge transfer. Analyses

of motivation were made on mean values for interest, stress, pressure/tension, perceived benefit, and perceived competence.

The overall effect of the teaching sequence was analysed by ANOVA, and differences between pre-test, post-test 1 and post-test 2 were examined using Tukey HSD.

## Result

### Effects on knowledge

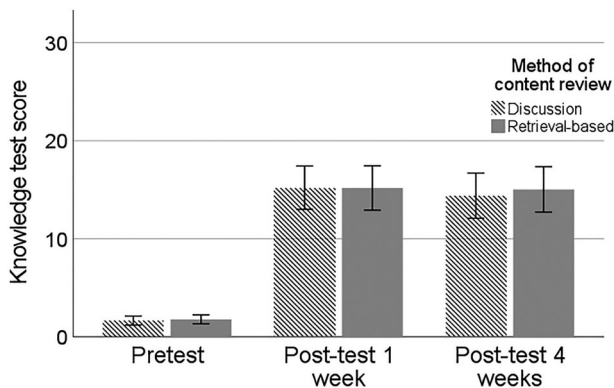
There was no difference in the pre-test in background knowledge between the group reviewing with retrieval-based learning and the group reviewing with discussion ( $p = 0.65$ ). Additionally, the pre-test mean score was very low, with a maximum possible score of 30, the mean score was only 1.7 points. The teaching sequence increased the students' knowledge about domestication and animal welfare noticeably (Figure 2).

There were no significant effects of method of reviewing on knowledge in the post-tests, whether measured as total score or as scores of simple or complex content, after one week (total score  $p = 0.87$ ; simple content  $p = 0.45$ ; complex content  $p = 0.86$ ) or after four weeks (total score  $p = 0.67$ ; simple content  $p = 0.15$ ; complex content  $p = 0.63$ ). No effects were found on ability of knowledge transfer after one week ( $p = 0.13$ ) or after four weeks ( $p = 0.67$ ).

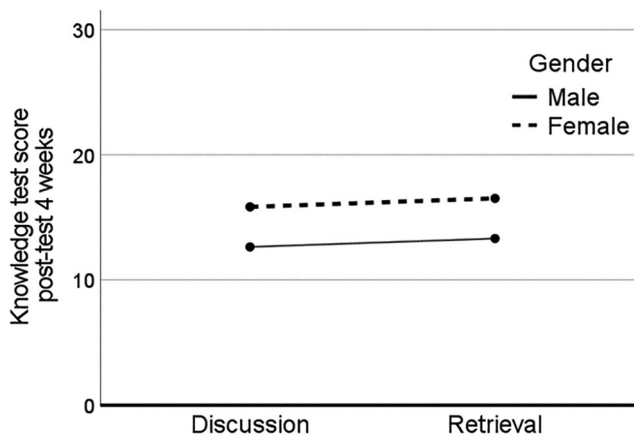
Moreover, the model showed no effects of school or gender, except for knowledge measured as total score after four weeks, where there was an effect of gender. Girls scored higher than boys regardless of method of reviewing (Figure 3).

### Effects on motivation

Significant effects of both method and school were observed for interest/enjoyment and perceived benefit of the content review among students (Table 1). Students in the



**Figure 2.** The instructional sequence, which included teaching at a zoological garden and two sessions of content review, either through retrieval-based learning or through discussion, was found to enhance students' knowledge of domestication and animal welfare ( $F = 137.1$ ,  $df = 2$ ,  $p < 0.001$ ). Prior to the instructional sequence, students obtained lower scores on the pre-test compared to both the post-tests conducted one week and four weeks after the final reviewing of content (Tukey HSD  $p > 0.001$ ). No statistical difference was detected between the post-tests (Tukey HSD  $p = 0.86$ ). Error Bars denotes 95% CI.



**Figure 3.** Girls scored higher than boys in the knowledge-test four weeks after the last content review, regardless of content reviewing method ( $F = 3.99$ ,  $df = 1$ ,  $p = 0.050$ ).

retrieval-based learning group reported higher levels of interest/enjoyment and experienced greater benefit compared to those in the discussion group, and interest/enjoyment was generally higher in school 1 than in school 2, regardless of the method of reviewing employed (Figure 4(A) and (B)). Regarding perceived competence, a school by method interaction was observed (Table 1), indicating that students in school 2 reported higher perceived competence in the retrieval-based group than in the discussion group, whereas no significant difference was found between groups in school 1 (Figure 4(C)). A significant gender by method interaction was observed for curiosity and the acquisition of new questions (Table 1), indicating that boys exhibited higher levels of curiosity and generated more new questions when reviewing through discussion, whereas girls benefited more from retrieval-based learning (Figure 4(D)). This interaction between gender and method of reviewing was apparent in both schools, and perceived curiosity was overall higher in school 1, regardless of method (Figure 4(D)). When testing the effect of method within gender and school no significant relationships were found, probably due to loss of statistical power due to reduced number of datapoints. No significant effects of method of reviewing, school or gender were observed for perceived pressure/tension (Table 1).

## Discussion

In summary the present study examined the use of two methods of reviewing content, group discussion and retrieval-based learning, in primary school classroom. Results indicated that both methods yielded comparable long-term knowledge retention, with no significant differences observed for simple and complex content or knowledge transfer. However, girls outperformed boys on the knowledge test administered four weeks after the last reviewing. With regards to motivation, students in the retrieval-based learning group reported higher levels of interest and perceived benefit than those in the discussion group. In terms of curiosity and new questions, a gender-by-method interaction was observed, revealing that girls in the retrieval group experienced greater curiosity and

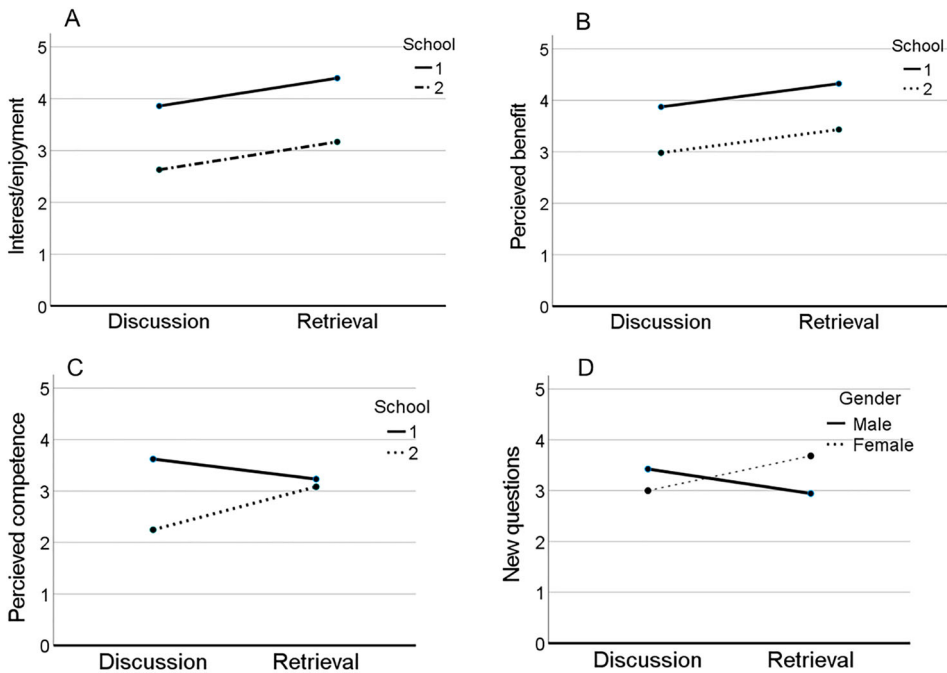
**Table 1.** Results from general linear analyses of method of reviewing content, discussion or retrieval-based, on different aspects of motivation.

Response variable	Source of variation	df	Mean	F	P
Interest/enjoyment	Method of reviewing	1	5.33	4.83	0.031
	Gender	1	1.8	1.64	0.205
	School	1	25.86	23.44	< 0.001
Perceived benefit	Method of reviewing	1	3.78	5.60	0.02
	Gender	1	0.48	0.71	0.40
	School	1	13.61	20.18	< 0.001
Perceived competence	Method of reviewing	1	0.85	1.32	0.78
	Gender	1	0.04	0.38	0.85
	School	1	9.93	1.55	0.43
	Method of reviewing * School	1	6.44	6.06	0.02
Pressure/tension	Method of reviewing	1	1.38	0.43	0.51
	Gender	1	0.37	1.16	0.29
	School	1	0.01	0.002	0.96
Use method more often	Method of reviewing	1	0.76	0.51	0.48
	Gender	1	4.02	2.69	0.11
	School	1	18.64	12.49	<0.001
Curiosity and new questions	Method of reviewing	1	0.19	1.30	0.72
	Gender	1	0.45	0.32	0.58
	School	1	5.59	9.60	0.01
	Method of reviewing *gender	1	6.27	4.33	0.04

more new questions than girls in the discussion group, whereas the opposite trend was indicated for boys. There were also several effects of school on motivation.

### Knowledge

Our study contributes to the rather limited collection of studies on younger students that have compared retrieval-based learning with another student-active teaching method in the classroom. Goossens et al. (2016), discovered that active recall improved primary school students' vocabulary acquisition more than generating words from a text, and Rowley and McCrudden (2020) noted that middle school students benefited more from active recall than writing notes when learning the content of a science text. In contrast to these studies, our findings did not indicate a significant effect of retrieval-based learning on knowledge retention. One commonly proposed explanation for the failure of retrieval-based learning to benefit children is that they may be unable to retrieve relevant content during practice (Karpicke et al., 2014). We do not regard this as a probable explanation for the absence of an effect in the current study, since students reviewing content through the retrieval-based method received clues to enhance their retrieval ability and were provided answers and information of correct content. An alternative explanation could be that the test-effect diminishes or vanishes as the complexity of the study content rises, as proposed by van Gog and Sweller (2015). Nevertheless, our examination of knowledge segmented into simple or complex content failed to reveal any distinction between the methods.



**Figure 4.** (A) Interest/enjoyment and (B) Perceived benefit were higher in the retrieval-based learning group compared to the discussion group, and overall higher in school 1 regardless of method. (C) Competence was perceived as higher in the retrieval-based group than in the discussion group in school 2 (ANOVA  $F = 7.6$ ,  $df = 1$ ,  $p = 0.008$ ) and there was no difference between groups in school 1 (ANOVA  $F = 1.03$ ,  $df = 1$ ;  $p = 0.32$ ). (D) A significant interaction of gender by treatment indicate that boys exhibited as greater level of curiosity and generated more new questions when engaging in discussion-based reviewing, whereas girls demonstrated these tendencies more prominently when utilising retrieval-based learning for review.

In our study, we initially taught the content by observations of real-life animals at the zoo and by conducting discussion and conversations that connected the content to the students' prior knowledge. This setting provided meaningful experiences; factors typically considered to facilitate learning achievement (Hattie, 2009). In contrast, Rowley and McCrudden (2020) and Goossens et al. (2016), utilised independent reading and explanations provided by a tutor for the students' initial exposure to the content in their respective studies. Therefore, a possible explanation for the divergent outcomes in these studies as compared to ours is that when the initial instructional setting strongly favours encoding, the effect of retrieval-based learning may be less pronounced. Instead, it could be the initial teaching setting, or a combination of the initial teaching setting and the comparative method used (discussion), that creates a situation in which retrieval-based learning is not more advantageous to the students. The effectiveness of retrieval-based learning appears to vary depending on the teaching method employed, as indicated by a classroom study conducted by Ortega-Tudela et al. (2021) that examined the use of retrieval practice in three different teaching situations with different content and teachers. They found a positive impact of retrieval practice on learning outcomes in two

out of three teaching situations, suggesting that the efficacy of this approach is contingent upon specific instructional contexts.

Finally, our only significant outcome on knowledge, that girls scored higher than boys on the knowledge test four weeks after the last repetition regardless of the repetition method, agrees with several findings that girls generally perform better than boys in the Swedish school system (Schleicher, 2019).

### **Motivation**

The importance of commitment and motivation as prerequisites for students' learning has been documented in previous research (Bybee et al., 2006; Deci & Ryan, 1985). Students' experience of retrieval-based learning is of particular interest, as previous studies on adults and adolescents have shown that the high cognitive effort required by this method can create a feeling of not mastering the content and cause students to underestimate its effectiveness (Karpicke et al., 2009; Karpicke & Grimaldi, 2012). To our knowledge, this study is the first to assess how the use of retrieval-based learning affects motivation of younger students. In contrast to the previous studies on adults, students in the retrieval-based learning group perceived the method as more beneficial for learning than those in the discussion group. Additionally, students in the retrieval-based method reported higher levels of interest than those in the discussion group, and we found no effects on pressure or tension. These findings suggest that the motivation of students who use retrieval-based learning may be higher than that of those who use discussion, in upper primary school settings. The underlying reasons for this could be that, in general, variation in methods is appreciated by learners (Blomgren, 2016) and that working with retrieval was something new to the students in our study.

In our study, a gender difference was observed in curiosity and the generation of new questions, suggesting that girls benefit more from retrieval-based learning while boys tend to benefit more from discussion. One possible explanation for this could be that girls are given less talking space in the science classroom (Eliasson et al., 2016), and therefore generate more new questions and curiosity when allowed to think undisturbed. Another explanation may be that girls this age are, on average, more capable of working independently, whereas earlier studies suggest that boys at this age require more immediate feed-back (Jones & Wheatley, 1990). Regarding perceived competence, there was school-by-method interaction, where students in the retrieval-based learning group in one school reported higher levels of perceived competence, whereas in the other school the relationship was absent. Furthermore, the school had an effect on all assessments of motivation except pressure/tension. Students in school number one perceived more interest, benefit, curiosity and wanted to use the method more often than students in school number two, regardless of the reviewing method. In this study we were not interested in investigating school differences. The inclusion of the school variable in the analysis is solely due to the participants being sourced from two distinct schools, which was necessary to increase the number of participants. When the dataset encompasses more than a single school, the school variable should be incorporated as a random factor in statistical analysis. This is not driven by any specific emphasis on differences among schools, but rather to account for potential variance that might exist unbeknownst to us. Therefore, we do not discuss these differences further, but it

can be mentioned that the schools were rather similar in respect to socioeconomic catchment area, school size, number of teachers in class and class size.

### **Limitations**

One limitation of the present study is the possibility that students need to practice the retrieval-based learning method before benefiting from the test-effect. In our study, discussion was a familiar approach for reviewing content among the students, while reviewing content with retrieval-based learning – where each student is given time to think independently in silence – was not commonly used. We observed that in the retrieval-based learning group several students immediately raised their hands and wanted to answer the questions, suggesting that the eagerness to communicate answers to others in the classroom may have hindered their ability to take the necessary time to make the cognitive effort and remember information.

Another limitation is the small-scale nature of the experiment, involving only two schools in Sweden and 81 students, more studies in more countries would be needed to generalise the result. Especially studies in non-western countries since cultural heritage may affect students' interest and learning in science (Ainley & Ainley, 2011).

### **Implications for science teaching and further studies**

Retrieval-based learning correlated with elevated levels of perceived benefit and interest among students, while also notably enhancing students' comprehension of domestication and animal welfare as long as four weeks after the last reviewing of content. Moreover, despite being a test-like situation students reported very low on levels of pressure and tension, and this did not differ from the experienced levels in the discussion group. This suggests that it holds promise as a valuable method in science education. In our research, we applied retrieval-based learning to review content initially imparted during an outdoor teaching experience at a zoo. Consequently, the method has the potential to be an advantageous strategy when integrated with other science teaching methods, such as outdoor education (e.g. Wünschmann et al., 2017), inquiry-based learning (e.g. Akaygun & Adadan, 2021), or engagement in scientific discourse (as reviewed by Bae et al., 2021).

Retrieval-based learning was not a prevalent approach for reviewing scientific content in the study's classroom, while the comparative method, namely discussion, enjoyed widespread use. Should this be a typical scenario in other primary science classrooms, employing this approach could aid teachers in attaining the goal of providing students with a learning environment enriched by diverse methods, as stipulated by the Swedish elementary school curriculum (Skolverket, 2022). Employing a diverse range of methods may heighten students' motivation, as noted by Blomgren (2016), and a systematic review by Nordenbo et al. (2008) showed that a variety of teaching methods contributes to student learning. Furthermore, utilising a variety of instructional approaches can promote an equitable learning environment in the classroom. Notably, in our study girls and boys responded differently to the employed teaching methods in terms of perceived level of curiosity, girls reporting higher curiosity when reviewing with retrieval-based learning. Thus, employing retrieval-based learning in the science classroom



could potentially enhance the educational experience for girls, which has reported to be given less talking space in the science classroom (Eliasson et al., 2016).

Moreover, it can be beneficial for students to learn the method of retrieval-based learning at a young age since is one of many powerful tools for self-study. Nonetheless, it is imperative for teachers to provide a metacognitive perspective by explaining why they are using this approach. This is important because, despite the largely positive experiences reported in our study, prior research has indicated that students might often underestimate the effectiveness of this method (Karpicke et al., 2009; Karpicke & Grimaldi, 2012). Probably due to the high degree of cognitive effort and that students may form a false connection between the ease of processing content and the degree to which they have learnt it (Karpicke et al., 2009).

Further studies should strive not only to be carried out in genuine classroom environments but also to offer ample opportunities for students to engage in retrieval-based learning and become better acquainted with the method before comparing it to alternative active repetition methods. Additionally, incorporating a teacher-implemented approach, such as that employed by Rowley and McCrudden (2020), may prove advantageous, as opposed to relying on the researcher-implemented retrieval practice that has been more commonly utilised in our and in prior investigations. Additionally, it would be beneficial for future studies to consider students' level of knowledge after the initial teaching event but before repetition. Although one study indicated that individual differences in personal traits and working memory capacity did not moderate the effect of the testing effect (Bertilsson et al., 2021), other studies on adolescents have demonstrated that for students with higher cognitive skills the method of studying is of less importance, whereas those with lower cognitive skills tend to benefit more from retrieval-based learning (e.g. Brewer & Unsworth, 2012). This finding is especially relevant in lower grades, where students' cognitive abilities may vary significantly. By analysing knowledge outcomes based on the initial teaching event, studies could investigate whether some students in primary science benefit more from retrieval-based learning than other methods of reviewing content.

## Acknowledgements

We are very grateful to those students and teachers who made this study possible, to Nordens Ark for use of the traditional farm facilities and to P. Börjesson for advising on statistical methods.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This study was supported by the remaining funds for the communication project "The time journey - human impact on the landscape through agriculture" to A. Jakobsson from the Swedish Council for Forestry and Agricultural Research under grant [017-00033], and by a grant from the institution of Department of Social and Behavioural studies at University West to A. Jakobsson and M. Kjörk.

## Ethics statement

The Ethical Review Board in Sweden has no ethical objections to the research project, in a consultative opinion with reference number 2021-02040.

## ORCID

Anna Jakobsson  <http://orcid.org/0000-0002-9355-2095>

Jenny Loberg  <http://orcid.org/0000-0002-4331-1998>

Maria Kjörk  <http://orcid.org/0009-0002-1618-9568>

## References

- Agarwal, P. K., Bain, P. M., & Chamberlain, R. W. (2012). The value of applied research: Retrieval practice improves classroom learning and recommendations from a teacher, a principal, and a scientist. *Educational Psychology Review*, 24(3), 437–448. <https://doi.org/10.1007/s10648-012-9210-2>
- Ainley, M., & Ainley, J. (2011). A cultural perspective on the structure of student interest in science. *International Journal of Science Education*, 33(1), 51–71. <https://doi.org/10.1080/09500693.2011.518640>
- Akaygun, S., & Adadan, E. (2021). Fostering senior primary school students' understanding of climate change in an inquiry-based learning environment. *Education 3-13*, 49(3), 330–343. <https://doi.org/10.1080/03004279.2020.1854961>
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. Holt, Rinehart, and Winston.
- Bae, C. L., Mills, D. C., Zhang, F., Sealy, M., Cabrera, L., & Sea, M. (2021). A systematic review of science discourse in K–12 urban classrooms in the United States: Accounting for individual, collective, and contextual factors. *Review of Educational Research*, 91(6), 831–877. <https://doi.org/10.3102/003465432111042415>
- Barton, C. (2018). *How I wish I had taught maths: Reflections on research, conversations with experts, and 12 years of mistakes*. John Catt Educational Ltd.
- Bertilsson, F., Stenlund, T., Wiklund-Hörnqvist, C., & Jonsson, B. (2021). Retrieval practice: Beneficial for all students or moderated by individual differences? *Psychology Learning & Teaching*, 20(1), 21–39. <https://doi.org/10.1177/1475725720973494>
- Bjork, E. L., & Bjork, R. A. (2014). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, & J. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 59–68). Worth.
- Bjork, R. A. (1975). Retrieval as a memory modifier. In R. Solo (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 123–144). Hillsdale.
- Bjork, R. A. (2011). On the symbiosis of learning, remembering, and forgetting. In I. A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 1–12). Psychology Press.
- Blomgren, J. (2016). *The elusive motivation: Students in a digitalized learning*, [Doctoral dissertation, Gothenburg University]. <http://hdl.handle.net/2077/47615>
- Brewer, G. A., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language*, 66(3), 407–415. <https://doi.org/10.1016/j.jml.2011.12.009>
- Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Carlson Powell, J., Westbrook, A., & Landes, N. (2006). *The BSCS 5E instructional model: Origins, effectiveness and applications*. Office of Science Education and National Institutes of Health.
- Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2), 109–134. [https://doi.org/10.1016/0092-6566\(85\)90023-6](https://doi.org/10.1016/0092-6566(85)90023-6)

- Didau, D. (2015). *What if everything you knew about education was wrong?* Crown House Publishing.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning With effective learning techniques. *Psychological Science in the Public Interest*, 14(1), 4–58. <https://doi.org/10.1177/1529100612453266>
- Earle, S., & McMahon, K. (2022, September 20). Applying cognitive science principles to primary science. *Impact*, 16. [https://my.chartered.college/impact\\_article/applying-cognitive-science-principles-to-primary-science/](https://my.chartered.college/impact_article/applying-cognitive-science-principles-to-primary-science/)
- Eliasson, N., Sørensen, H., & Karlsson, K. G. (2016). Teacher–student interaction in contemporary science classrooms: Is participation still a question of gender?†. *International Journal of Science Education*, 38(10), 1655–1672. <https://doi.org/10.1080/09500693.2016.1213457>
- Fazio, L. K., & Marsh, E. J. (2019). Retrieval-Based learning in children. *Current Directions in Psychological Science*, 28(2), 111–116. <https://doi.org/10.1177/0963721418806673>
- Fritz, C. O., Morris, P. E., Nolan, D., & Singleton, J. (2007). Expanding retrieval practice: An effective aid to preschool children's learning. *Quarterly Journal of Experimental Psychology*, 60(7), 991–1004. <https://doi.org/10.1080/17470210600823595>
- Goossens, N. A. M. C., Camp, G., Verkoeijen, P. P. J. L., Tabbers, H. K., Bouwmeester, S., & Zwaan, R. A. (2016). Distributed practice and retrieval practice in primary school vocabulary learning: A multi-classroom study. *Applied Cognitive Psychology*, 30(5), 700–712. <https://doi.org/10.1002/acp.3245>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Jaeger, A., Eisenkraemer, R. E., & Milnitsky Stein, L. (2015). Test-enhanced learning in third-grade children. *Educational Psychology*, 35(4), 513–521. <https://doi.org/10.1080/01443410.2014.963030>
- Jones, G. M., & Wheatley, J. (1990). Gender differences in teacher-student interactions in science classrooms. *Journal of Research in Science Teaching*, 27(19), 861–874. <https://doi.org/10.1002/tea.3660270906>
- Kang, S. H. K., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, 19(4/5), 528–558. <https://doi.org/10.1080/09541440601056620>
- Karpicke, J. D., Blunt, J. D., & Smith, M. A. (2016). Retrieval-based learning: Positive effects of retrieval practice in elementary school children. *Frontiers in Psychology*, 7, 1–9. <https://doi.org/10.3389/fpsyg.2016.00350>
- 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., & Sumeracki, M. A. (2014). Retrieval-based learning: The need for guided retrieval in elementary school children. *Journal of Applied Research in Memory and Cognition*, 3(3), 198–206. <https://doi.org/10.1037/h0101802>
- Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory (Hove, England)*, 17(4), 471–479. <https://doi.org/10.1080/09658210802647009>
- Karpicke, J. D., & Grimaldi, P. J. (2012). Retrieval-based learning: A perspective for enhancing meaningful learning. *Educational Psychology Review*, 24(3), 401–418. <https://doi.org/10.1007/s10648-012-9202-2>
- Karpicke, J. D., & Zaromb, F. M. (2010). Retrieval mode distinguishes the testing effect from the generation effect. *Journal of Memory and Language*, 62(3), 227–239. <https://doi.org/10.1016/j.jml.2009.11.010>
- Lim, S. W. H., Ng, G. J. P., & Wong, G. Q. H. (2015). Learning psychological research and statistical concepts using retrieval-based practice. *Frontiers in Psychology*, 6, 1484. <https://doi.org/10.3389/fpsyg.2015.01484>
- 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>

- Määttä, K., & Uusiautti, S. (2020). Nine contradictory observations about girls' and boys' upbringing and education – the strength-based approach as the way to eliminate the gender gap. *Frontiers in Education*, 5(134), <https://doi.org/10.3389/feduc.2020.00134>
- McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger, H. I. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology*, 103(2), 399–414. <https://doi.org/10.1037/a0021782>
- McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., & Roediger, H. I. (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>
- Mellor, D., & Moore, K. (2014). The use of Likert scales with children. *Journal of Pediatric Psychology*, 39(3), 369–379. <https://doi.org/10.1093/jpepsy/jst079>
- Moreira, B. F. T., Pinto, T. S. S., Starling, D. S. V., & Jaeger, A. (2019). Retrieval practice in classroom settings: A review of applied research. *Frontiers in Education*, 4(5), 1–16. <https://doi.org/10.3389/feduc.2019.00005>
- Nordenbo, S. E., M, S. L., Tiftikçi, N., Wendt, R. E., & Østergaard, S. (2008). *Teacher competences and pupil learning in pre-school and school – A systematic review carried out for the Ministry of Education and Research*. Danish Clearinghouse for Educational research.
- Ortega-Tudela, J. M., Lechuga, M. T., Bermúdez-Sierra, M., & Gómez-Ariza, C. J. (2021). Testing the effectiveness of retrieval-based learning in naturalistic school settings. *SAGE Open*, 11(4), 215824402110615–7. <https://doi.org/10.1177/21582440211061569>
- Perkins, D. N., & Salomon, G. (1992). Transfer of learning. In T. Husen, & T. N. Postelwhite (Eds.), *The international encyclopedia of education* (Vol. 11, 2 ed, pp. 6452–6457). Pergamon Press.
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20–27. <https://doi.org/10.1016/j.tics.2010.09.003>
- Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181–210. <https://doi.org/10.1111/j.1745-6916.2006.00012.x>
- Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(1), 233–239. <https://doi.org/10.1037/a0017678>
- 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>
- Sanders, L. L. O., Ponte, R. P., Júnior, A. B. V., Junior, A. A. P., & Furtado Leitão, M. K. A. M. (2019). Retrieval-based learning in neuroanatomy classes. *REVISTA BRASILEIRA DE EDUCAÇÃO MÉDICA*, 43(43), 92–98. <https://doi.org/10.1590/1981-52712015v43n4rb20180184ingles>
- Schleicher, A. (2019). *Pisa 2018. Insights and interpretations*. OECD Publishing.
- Schmidt, H. G., & Mamede, S. (2020). How cognitive psychology changed the face of medical education research. *Advances in Health Sciences Education*, 25(5), 1025–1043. <https://doi.org/10.1007/s10459-020-10011-0>
- Skolverket. (2022). *Curriculum for the compulsory school, preschool class and school-age educare 2022*. The Swedish National Agency for Education.
- Taherdoost, H. (2019). What Is the best response scale for survey and questionnaire design? Review of different lengths of rating scale/attitude scale/Likert scale. *International Journal of Academic Researching Management*, 8(1), 9–10.
- Toli, G., & Kallery, M. (2021). Enhancing student interest to promote learning in science: The case of the concept of energy. *Education Sciences*, 11. <https://doi.org/10.3390/educsci11050220>
- Underwood, A. J. (1997). *Experiments in ecology: Their logical design and interpretation using analysis of variance*. Cambridge University Press.
- van Gog, T., & Sweller, J. (2015). Not new, but nearly forgotten: The testing effect decreases or even disappears as the complexity of learning materials increases. *Educational Psychology Review*, 27(2), 247–264. <https://doi.org/10.1007/s10648-015-9310-x>

- Wiklund-Hörnqvist, C. (2014). Brain-based teaching. Behavioral and neuro-cognitive evidence for the power of test-enhanced learning [doctoral dissertation]. Umeå. University].
- Wilde, M., Bätz, K., Kovaleva, A., & Urhahne, D. (2009). Überprüfung einer kurzskala intrinsischer motivation (testing a short scale of intrinsic motivation). (*KIM*). *Zeitschrift für Didaktik der Naturwissenschaften (Journal of Science Teaching Methods)*, 15, 31–34.
- Wünschmann, S., Wüst-Ackermann, P., Randler, C., Christian Vollmer, C., & Itzek-Greulich, H. (2017). Learning achievement and motivation in an out-of-school setting—visiting amphibians and reptiles in a zoo is more effective than a lesson at school. *Research in Science Education*, 47 (3), 497–518. <https://doi.org/10.1007/s11165-016-9513-2>

## Appendices

### Appendix 1

Questions in the knowledge test that were administered to students one week and four weeks after their last repetition. The questions are translated from Swedish. The questions were classified into simple or complex content, and this classification is denoted in parentheses after the question

1. How does breeding take place? (not classified)
2. What does "to domesticate" mean? (not classified)
3. What is the difference between to tame and to domesticate? Mark ONE of the options you think is right. (complex)
  - To tame means to have accustomed an individual animal to humans. To domesticate means that many people have influenced a group of animals over a long period of time."
  - Domesticate means to accustom an individual animal to humans. Taming means that many people have influenced a group of animals for a long time.
  - Taming is more heavy-handed. For example, a whip is often used.
  - Domesticate is more heavy-handed. For example, a whip is often used.
  - The difference is that taming is about predators and domestication is about herbivores.
  - The difference is that domestication is about predators and taming is about herbivores.
4. Even though dangerous animals cannot enter the hen house, the hen still feels safest if it is allowed to perch above the ground. Why? (complex)
5. Why is it that the female of the chicken's wild relative is always grey-brown, while both male and female of our domesticated chickens can be brightly coloured? (complex)
6. Write the name of a breed of chicken that is particularly good at laying eggs. (simple)
7. What is the scientific name for hen? (simple)
8. Name a country where the chicken's wild relative lives. (simple)
9. What can happen to a hen that does not have access to sand? Tick ONE answer for what you think. (simple)
  - The hen may get hungry because there is a lot of food in the sand.
  - The hen may get cold because hens usually warm themselves in sand pits.
  - The hen can get dirty because she needs sand to keep herself clean.
  - The hen may get tired because hens like to rest in sand pits.

- The hen may get bored if she is not allowed to play in the sand.
10. What is the name of the wild ancestor of the domesticated chicken? (simple)
11. Why do all piglets have their own teat when nursing from the mother pig? Tick ONE answer for what you think. (complex)
- The milk in the teats tastes different, and the piglet gets used to a particular taste.
  - The wild relative of the pig needed to keep order with her piglets so that the young of other mothers did not come and steal milk.
  - Piglets who are extra hungry have been butting their teat to get more milk.
  - So that the piglets don't infect each other.
  - To make it fair so that everyone gets the same amount of milk.
12. Even though the pigs get as much food as they need from the farmer, they still want to root in the soil. What is the most important reason? Tick ONE answer for what you think. (complex)
- It's nice because the pigs' snouts often itch.
  - Since their wild relatives searched for food in this way, rooting has become a habit that pigs still have.
  - There is particularly good food in the soil, such as earthworms, which are like candy for pigs.
  - There is particularly good food in the soil, such as earthworms, which are like candy for pigs.
  - They need to grind their teeth. The teeth become too long if they are not allowed to root.
13. What is the scientific name for pig? (simple)
14. The male pig is called a boar. What is the female called? (simple)
15. What does the female pig want to do before she gives birth to her young? (not classified)
16. In some countries, pigs' tails are cut off. What do we do in Sweden to avoid having to do that? (complex)
17. Why do horses need to stay in paddocks that are quite large? Tick ONE answer for what you think. (simple)
- Because each horse wants its own territory.
  - Because the leader stallion and the leader mare want quite a lot of space for themselves.
  - Because its wild relative moved over large areas.
  - So that there is enough grass to eat.
  - Because it has an instinct to run far away when it gets scared, like when a dog comes too close to the fence.
18. The wild relative to the horse does not exist any longer. Why is that? (simple)
19. What is the scientific name of the horse? (simple)
20. What was the reason why man first started having the horse as a farm animal? (simple)
21. Horses eat a large part of their waking hours. What happens if the horse is not allowed to do that and why is that? (not classified)

This will happen:

This is because:

22. There are many different breeds of rabbits that look very different. How did they come about? (transfer) (The students were here provided pictures of a lop rabbit, an angora rabbit, and a pygmy rabbit.)

## Appendix 2

### *Multiple-choice questionnaire about motivation*

Presented below are the questionnaire taken by the students after the session of reviewing content at school. The questionnaire is translated from Swedish. The questions are first shown grouped under motivation aspect, and presented as given to the students in the questionnaire .

#### **Interest/enjoyment**

1. I found the reviewing fun to do.
2. I found the reviewing interesting.
3. I found the reviewing entertaining.

#### **Perceived competence**

1. I am pleased with my prestaton during reviewing.
2. During reviewing I felt smart.
3. I think I was rather good at answering the questions during reviewing.

#### **Perceived pressure/tension**

1. I felt pressure during reviewing.
2. I felt tense during reviewing.
3. I was worrying about not doing good enough.

#### **Perceived benefit**

1. I have learnt a lot during reviewing.
2. I think I will be able to answer more questions next time.
3. I feel that it is good to do reviewing the way we did.

#### **Singel questions**

1. I would like to use this method of reviewing more often in school.
2. During reviewing I have gained many new questions that I am curious about.

Below, the questions are presented in the order they appear in the questionnaire. The Likert scale is only displayed here for the first question.

1. I found the reviewing fun to do.

Not at all true	A bit true	Partially true	Pretty much true	Very true
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

1. I felt pressure during reviewing.
2. I have learnt a lot during reviewing.
3. I found the reviewing entertaining.
4. I felt tense during reviewing.
5. During reviewing I felt smart.
6. I found the reviewing entertaining.
7. I think I will be able to answer more questions next time.
8. I was worrying about not doing good enough.
9. I think I was rather good at answering the questions during reviewing.
10. I feel that it is good to do reviewing the way we did.
11. I am pleased with my prestation during reviewing.
12. I would like to use this way of practicing more often in school.
13. I have received many new questions that I wonder about and am curious about.