Psicologia Escolar e Educacional

RETRIEVAL PRACTICE: WHICH FACTORS SHOULD EDUCATORS PAY ATTENTION TO?

Roberta Ekuni [1]; Sabine Pompeia [2]

ABSTRACT

Retrieval practice, or the testing effect, is a study technique that involves trying to remember information to which we were previously exposed. Although this practice increases the long-term-retention of information compared to traditional study techniques, among several other advantages with ample scientific evidence, this strategy is not usually the most often used among students. Educators should help students to use this technique in their daily lives. In order to optimize its applicability, this article discusses which factors interfere in this practice, including: the importance of feedback, the way in which retrieval practice is carried out and the response format, the number of repetitions of retrieval attempts to recall information and the interval between these repetitions. The appropriation of knowledge about these factors positively influences the implementation of this technique in the classroom, thus promoting evidence-based education.

Keywords: cognitive processes; learning; memory.

Práctica de recordar: ¿a cuáles factores los educadores deben atentarse?

RESUMEN

La práctica de recordar (*retrieval practice*) el efecto de la verificación (*testing-effect*) es una estrategia de estudio que abarca tentar lembrar informaciones a las cuales fuimos anteriormente expuestos. Aunque esa práctica aumente el tiempo de retención de informaciones comparada a las formas tradicionales de estudiar, entre diversas otras ventajas con amplia evidencia científica, esa estrategia no costumbra ser la más usada entre alumnos. Educadores deben, así, auxiliar estudiantes a utilizar esa estrategia en su cotidiano. Con el intuito de optimizar su aplicabilidad, en el presente artículo discute cuales factores interfieren en esa práctica, incluyendo: la importancia de *feedback*, la forma con que la práctica de recordar es realizada y el formato de respuesta de los alumnos, el número de repeticiones de tentativas de recordar informaciones y el intervalo entre esas repeticiones. La apropiación del conocimiento sobre esos factores influencia positivamente la implantación de la técnica en sala de clase, promoviendo así una educación basada en evidencias.

Palabras clave: procesos cognitivos; aprendizaje; memoria.

Prática de lembrar: a quais fatores os educadores devem se atentar?

RESUMO

A prática de lembrar (*retrieval practice*) ou efeito da testagem (*testing-effect*) é uma estratégia de estudo que envolve tentar lembrar informações às quais fomos anteriormente expostos. Embora essa prática aumente o tempo de retenção de informações comparada às formas tradicionais de estudar, dentre várias outras vantagens com ampla evidência científica, essa estratégia não costuma ser a mais usada entre alunos. Educadores devem, assim, auxiliar estudantes a utilizarem essa estratégia em seu cotidiano. Com o intuito de otimizar sua aplicabilidade, o presente artigo discute quais fatores interferem nessa prática, incluindo: a importância de *feedback*, a forma com que a prática de lembrar é realizada e o formato de resposta dos alunos, o número de repetições de tentativas de recordar informações e o intervalos entre essas repetições. A apropriação do conhecimento sobre esses fatores influencia positivamente a implantação da técnica em sala de aula, promovendo assim uma educação baseada em evidências.

Palavras-chave: processos cognitivos; aprendizagem; memória.

[1] Universidade Estadual do Norte do Paraná – Bandeirantes – PR – Brasil; robertaekuni@gmail.com

[2] Universidade Federal de São Paulo – São Paulo – SP – Brasil; spompeia@gmail.com

INTRODUCTION

Recent data indicate that Brazil occupies one of the worst positions in the education ranking (UNESCO, 2014). This contrasts sharply with Brazilian scientific production in this area: ninth position in scientific production in the World and first place in Latin America (SCImago, 2019). This contradiction may hint at a disconnection between scientific production in this field and its use in the classroom. The result is an education system that is not based on scientific knowledge and does not promote internationally recommended evidence-based teaching practices (see Bruer, 1997; Stern, 2005).

The use of evidence-based educational practices relies on scientific findings rather than on opinions, empiricism and/or tradition (Slavin, 2020). Evidence demonstrates that retrieval practice, which consists of trying to remember information to which one was previously exposed, is a study strategy that has a number of advantages. This practice, also known as the testing-effect, consists of doing exercises, answering questions, and/or taking tests (Roediger & Karpicke, 2006b), asking questions about the material, associating it with other contents, etc. In addition to promoting long-lasting retention of information (Karpicke & Roediger, 2007; Roediger & Karpicke, 2006a; Rowland, 2014), this way of studying: (a) allows people to find gaps in their knowledge, which encourages them to study other related topics; b) promotes a better mental organization of the learned material; (c) improves the transfer of knowledge to other contexts; d) assists in the memorization of content associated with what one tried to remember; e) reduces confusion between similar materials (Roediger, Putnam, & Smith, 2011). There are many theories that seek to explain how this way of studying improves learning (Rowland, 2014; van den Broek et al., 2016), a subject that will not be discussed in detail here, as this is beyond the scope of this paper.

Unsurprisingly, it is clear that people learn more content if they practice retrieving information compared to not studying at all. However, studies also show that retrieval practice is more efficient than other ways of studying. In most published papers, the control condition for comparison consists of rereading the material. That is, researchers compare how much information people retain when they study using retrieval practice (such as answering tests) versus when they re-read the same material that was tested for in another condition. When there are, for example, three testing opportunities regarding a certain to-be-learned material, the retention of this information is compared to a condition in which the same material is reread three times. Studies that focused on classroom activities, academic content with a variety of disciplines and a diversity of ages, show that retrieval practice is more efficient than rereading the material (e.g. Agarwal, D'Antonio, Roediger, McDermott, & McDaniel, 2014; McDaniel, Agarwal, Huelser, McDermott, & Roediger, 2011).

Despite the abundant scientific evidence corroborating that trying to recall content (retrieval practice) is good for learning, students commonly prefer less effective study strategies such as rereading notes or going over the course materials before exams (Karpicke, Butler, & Roediger, 2009). When engaging in rereading , the connections within the studied content are reinforced as presented in the written material. Hence, the activation of knowledge bases (semantic networks) occur in similar ways (e.g. following the same sequence of information associations). This way, information is not well integrated, nor well associated with other areas of knowledge, because only specific associations are reinforced. Thus, if students need to use the information in other contexts, or attempt to recall pieced of the studied information with questions that do not reflect how the information was actually studied, they will not be able to do so efficiently. The good news is that it is possible to improve students performance with the dissemination and implementation of the use of retrieval practice as a means of studying. This is easy to implement, does not necessarily involve extra time in class or study time, nor does it require additional financial resources (Roediger & Pyc, 2012b), as will be detailed below.

In order to optimize the applicability of retrieval practice, educators need to understand which factors increase the effectiveness of this technique. Therefore, this article aims to discuss factors that are known to increase the positive effects of retrieval practice in order to help educators to efficiently implement this technique.

In order to understand what will be explained below it is important to explain how scientific experiments on retrieval practice are designed. In general, experiments in the lab or in real classroom environments include several steps. The first step is exposure to the to-be-learned content in a format. This is generally similar to the forms traditionally used in the classroom, such as reading a text, learning information on lists, or attending a lesson given by a teacher.

The second step involves retrieval practice itself, as a rule by means of tests, exercises or questions related to the previously exposed content. This can be done in different formats, such as multiple-choice or open-answer tests. For this reason, in this literature, retrieval practice is usually also referred to as the "testing *effect".* Tests, or retrieval practice opportunities, can be carried out only once or several times, at different time intervals, ranging from seconds to hours or days. Keep in mind, however, that retrieval practice does not

necessarily involve testing itself, as mentioned above. For example, summing up what was learned in class is a form of practicing retrieval that is not exactly a "test." Therefore, when the word "testing" or "test" is mentioned in this context, consider that they reflect retrieval practice opportunities by means of questions, exercises, or any form of stimulating students to try to remember academic content of any kind.

The last step involves assessment of the retention of information after a variable length of time since the last retrieval practice opportunity, or testing. This assessment is done by means of an exams or finals and aims to evaluate the extent to which retrieval practice influences the retention of information in the long term.

The analysis of the role of factors that influence the effectiveness of retrieval practice in this type of experiment is not simple because studies vary widely in terms of many aspects. For this reason, here we explore the factors that: a) can be employed in educational contexts and that are most likely to result in long-lasting learning; and b) involve as little effort as possible on the part of educators in terms of time spent ministering content, and applying and correcting assignments.

Feedback

Feedback, which consists of informing students whether the answers given in retrieval practice opportunities were correct or not, increases the magnitude of the beneficial effects of retrieval practice (Pashler, Cepeda, Wixted, & Rohrer, 2005; Rowland, 2014). Feedback is particularly important when tests involve recognition, such as when using multiple-choice formats, because students are exposed to wrong alternatives that may appear to be correct. Thus, after taking a multiple-choice tests without feedback, students may inadvertently retain incorrect information and miss out on the opportunity to learn the right answer (Roediger & Butler, 2011). Feedback does not necessarily have to be given by the teacher. The test answers can be consulted on support material. The most important thing to bear in mind is that students must have the opportunity to check which was the correct answers.

Feedback has a beneficial effect regardless of whether or not the answer was correctly remembered or not (Butler, Karpicke, & Roediger, 2008). What matters is that the student try to remember and then have access to the answer. When the student does not know the answer, or has low confidence that he/she got it right, feedback has a greater positive effect (Butler et al., 2008) than when students have high confidence that they got a test question right (Hays, Kornell, & Bjork, 2010). Since it is not possible to know students' degree of confidence in advance in real classroom environments, it is recommended

that there should always be feedback. Being informed about the correct answer has other benefits, such as stabilizing marginal knowledge, that is, the knowledge that people can recognize, but not recall without clues (Cantor, Eslick, Marsh, Bjork, & Bjork, 2014). In addition, there is evidence that feedback increases semantic activation related to the provided answers (Hays et al., 2012), that is, it also promotes the consolidation of information *associated* with the material being learned. Furthermore, feedback also improves metacognition, which is the ability to detect how much of the presented content has been retained (Butler et al., 2008).

There is controversy about the advantages and disadvantages of the time lag between retrieval practice attempts and feedback. Some consider that postponing feedback until a long time after retrieval practice does not allow students to take advantage of the activation of semantic networks triggered by feedback (Hays et al., 2012). Others argue that it is wiser to provide feedback later rather than immediately because, this way, content can be reviewed in a distributed manner over time and not all at once. Delayed feedback may also confer advantages for memorization (Smith & Kimball, 2010). Additionally, providing feedback immediately after the answer is given and having students try to answer the questions until they find out the correct answer are equivalent practices in term of long-term retention (Butler, Karpicke, & Roediger, 2007). However, what is considered in these studies as "immediate" or "late" feedback varies widely in terms of time spans, so it is not yet clear what is the ideal lag.

The recommendation that seems to us to be more sensible is that teachers should always allow students to have access to answers whenever this is possible. Students should also have access to the answers they have given, so that they can correct answers they got wrong. We believe that, in the classroom, it is important not to postpone feedback for too long after the activity, because there is a risk that students will not even remember why they chose a given answer and not another. This reduces the possibility that they will actually integrate the right answers into their knowledge network.

Test format and characteristics

The types of tests mostly used to study retrieval practice are multiple-choice questions, in which the correct answer must be recognized among several alternatives (McDaniel, Roediger, & McDermott, 2007). Another type of test that also involves recognizing correct answers consists of true or false question formats (Roediger & Butler, 2011). Tests may also involve free recall, such as open-ended questions (Roediger et al., 2011), or cued recall, such as presenting the first letter of to-be-remembered words or terms (Cull, 2000).

According to Roediger, Agarwal, McDaniel, and McDermott (2011), open-ended questions are best suited to elicit reorganization (reconsolidation) of learned content. When compared to multiple-choice questions, open-ended tests proved to be better at promoting long-term retention (Kang, McDermott, & Roediger, 2007; McDaniel, Anderson, Derbish, & Morrisette, 2007; McDaniel, Roediger & McDermott, 2007; Rowland, 2014). However, there is a certain controversy over this issue (see McDermott, Agarwal, D'Antonio, Roediger, & McDaniel, 2014). These differences of opinion and results in the literature stem primarily from differences in the characteristics of the alternative answers in multiple-choice tests. Little, Bjork, Bjork, and Angello (2012) showed that multiple-choice tests that are well formulated lead to better information retention, that is, those that have plausible alternatives, which are not too easy or obvious. Tests such as these involve more effort in recognizing the correct answer than merely identifying the alternative that seems more reasonable as an answer. In other words, poorly created multiple-choice tests may not exactly involve retrieval practice, but might rather tap the use of logic to identify the answer which is most likely to be correct.

Therefore, the effectiveness of multiple choice tests in promoting long-lasting retention of information depends on the "quality" of the alternatives. Another issue that deserves attention is that when alternatives are very similar to each other they can lead students to choose the wrong alternative and persist in this answer when they try to remember the information later. This effect is called negative suggestion (Marsh, Agarwal, & Roediger, 2009; Roediger & Butler, 2011) and may result in mislearning the content. Fortunately, this can be minimized by providing feedback (Roediger & Marsh, 2005).

Another test format that is much less explored in the literature is fill-in-the-blank questions, in which the participant must recall one or a few keywords. This type of test has not been properly studied. There are suggestions that tasks like this produce poor retrieval practice effects (de Jonge, Tabbers, & Rikers, 2015) because they require little integration and elaboration of content (Karpicke & Aue, 2015).

Retrieval practice by means of tests provided by teachers or created by peers have similar beneficial effects (Weinstein, McDermott, & Roediger, 2010). However, generating and answering questions is more time-consuming for students (Weinstein et al., 2010). It is therefore recommended that tests be made available to students because of their limited time to study, especially because teachers know what is the essential content that deserves attention and should be tested for. However, trying to come up with questions is a good type of exercise that involves retrieval practice.

Therefore, it is a good way of studying that can additionally result in a question bank for the teacher to use in other suitable contexts.

Concerning the type and characteristics of tests, the message to educators is: try to stimulate retrieval practice in class or as homework by providing exercises or open-ended tests questions, always followed by feedback after giving students enough time to process information. Feedback should also be provided within a short enough time frame considering that students must still remember the answers they provided, and why they did so. The correction of open-ended written answers, however, is taxing for teachers in terms of time.

Multiple-choice tests also work well as ways to stimulate retrieval practice but present some disadvantages in terms of the difficulty of creating a balance between: 1) the selection of alternatives that actually challenge students to spot the right answer; and 2) avoiding alternatives that potentially lead students to associate questions with wrong answers, when many alternatives are too similar to each other. Creating multiple-choice tests with these characteristics also takes time, which teachers seldom have. Potential problems in selecting alternatives, however, can be minimized by providing feedback.

It is important to point out that feedback on all types of tests can be done collectively, which leads to the same benefits as individualised corrections (see Butler et al., 2007). Therefore, providing feedback is not necessarily too time consuming and the eventual loss of time is compensated for by better learning on the part of students.

Exam/final exam formats

Traditionally, tests can be of various formats, such as multiple-choice, open-ended answers, cued-recall, fill-in-the-blank, etc. All these formats can also be used in finals or exams, which assess the long-term retention of the course material. Hence, exams are perfect circumstances not only for evaluating students, but also for them to practice retrieval. This is so because exams involve total concentration without interruptions (such as from cell phones), something that students rarely experience while studying by themselves. In other words, well-formulated finals can also help students learn.

According to Rowland's meta-analysis (2014), the beneficial effects of retrieval practice can be observed even when there is no congruence between the test format used to practice retrieval and the final exam. For example, when retrieval practice involves open-answer questions and the finals involves multiple-choice tests. This is known as cross-format, i.e., the format of the tests differs from the one used in the final exam (see McDaniel, Anderson et al., 2007). The advantage of

using cross-format is twofold: 1) it allows teachers to assess the transfer of information to another context, thus avoiding the measurement of simple memorization of answers (see McDaniel, Anderson et al., 2007) that can happen if exactly the same question, with the same answer alternatives, are used during retrieval practice and again on the final test; and 2) it consists of another occasion for practicing retrieval in a diversified way because the questions are different, which makes the final test a new learning opportunity (feedback should be given, as explained earlier).

As to whether the test questions should be presented in an increasing order of difficulty, that is, starting with easier questions and ending with harder ones, it has been experimentally demonstrated that this does not affect performance. However, this influences the perception of students: under these conditions, they tend to have a rather optimistic view on their own performance (Weinstein & Roediger, 2010, 2012). This can be a problem if there is no feedback or students are not quickly warned about their grade. Without this, they may think they did so well that they will not have to study for a recovery test, for example.

Teachers should therefore ideally try to prepare exams/finals that have questions that are different from those to which students had access to when they were preparing for the exams. This allows the teacher to verify the transfer of retained content to other contexts (see McDaniel, Anderson et al., 2007), that is, truly measure how much students learned, and also stimulate learning during exams themselves. Feedback must always be provided, individually or collectively. This increases the probability of learning and provides the opportunity for students to realize how much they know of the subject, so that they can adjust their subsequent study scheme. If the option is to use tests with multiple-choice questions, consider the points mentioned above when it comes to formulating the alternative answers.

Characteristics of test response modalities

In relation to the type of response given during retrieval practice, there seem to be no differences whether they are written, typed, spoken aloud (overt retrieval) or only thought (covert retrieval) (Putnam & Roediger, 2012; Smith, Roediger, & Karpicke, 2013). When questions are asked in the classroom, in contrast, it is difficult to establishing whether students are actually thinking about possible answers. One way around this is the proposal in a Teacher's Guide on how to use Retrieval Practice.[1] The strategy involves providing students with a set of colored paper cards, each one for a specific multiple-choice question answer alternative. For example, a blue card for alternative A,

green for B, etc. The teacher writes the question and the alternative answers on the board or projects them onto a screen. Students can then be given time to think about the answer and be collectively prompted at a specific moment to lift the card with the answer they think is correct. Thus, there is a guarantee that everyone will work hard to respond (Agarwal, Roediger, McDaniel, & McDermott, 2013), not just those few students who usually do so. The cards are colored to help the teacher detect errors more easily. Another option is to ask questions during class, give students time to think about the answer, and then draw a student to say the answer aloud. As students do not know who will be picked, everyone will make an effort to think about the answer, which characterizes retrieval practice (Kang et al., 2007). This, however, can generate stress in the class, which is totally unnecessary and demotivating.

Another alternative, also from the Guide described above, is to ask students to laminate a sheet of cardboard so that it becomes an "erasable board", on which they can write answers with color markers and then erase them with alcohol or any other similar cleaner (Agarwal et al., 2013). To improve retention, the teacher should always provide the correct answer (feedback) (Rowland, 2014).

We believe that students should not be punished if they provide a wrong answer, or rewarded when they get it right. Remember that retrieval practice works in both cases if there is feedback. Therefore, students should be encouraged to practice retrieving information, regardless of whether or not the answers they come up with are the correct ones.

Another point of interest is that the benefit of practicing retrieval before an exam seems to be independent of whether or not consultation to other sources, such as the Internet, is allowed during the exam, as long as students do not know whether the exam will be open- or closed-book (Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008). When they know, sometimes there is a drop in performance (about 10% in Agarwal & Roediger, 2011). According to Agarwal and Roediger (2011), this happens because the study time tends to be shorter when students prepare for an open-book exam. That is, expectations regarding the possibility of consulting the course material influences how much students study and this can negatively impact their outcome. In order to avoid this effect, an appropriate orientation concerning the way of studying in the case of open-book tests should be: "studying helps you know answers and also to locate where information can be found in the support material in the case of open book exams ". If this is explained to students, open-book exams can be a good way for creating a practical opportunity to retrieve information, with the consequent benefits for learning.

---

[1] Available at http://www.retrievalpractice.org/baixe-o-guia. Accessed Feb 27, 2019.

An interesting point about exams is that students should always be encouraged to answer questions, take tests, or do exercises, even if they do not know the exact answer. Studies show that trying to provide a suitable response in these conditions improves students' performance in subsequent exams compared to leaving questions unanswered (Kornell, Hays, & Bjork, 2009). One way to do this is to give some kind of reward for answer attempts in order to encourage students to keep on trying (Ekuni & Pompeia, 2015).

Number of repetitions of retrieval practice attempts and the time interval between them

Stimulating students to try to recall information leads to better retention than asking them to reread texts or class notes (Rowland, 2014), and this occurs even if multiple testing is compared with multiple rereading regarding the same information (Karpicke & Roediger, 2007).

We know that the more we devote ourselves to specific content, the better our learning will be. However, in practice, students have a myriad of things to study. Thus, a frequent question is: how often do students have to practice retrieval, and at what intervals, in order to achieve a more efficient and long-lasting learning. Rowland (2014) shows that the biggest gain in retention is achieved by attempting to recall information for the first time. Retrieving the same content repeatedly increases retention even more, but the gains with each new retrieval attempt has diminished returns (Bangert-Drowns, Kulik, & Kulik, 1991; Rowland, 2014). It is unclear, however, what is the ideal number of repetitions of retrieval practice attempts regarding the same content. McDermott et al. (2014) suggest that two tests with feedback on actual classroom content are sufficient to ensure long-lasting learning, but these authors do not address the issue of time intervals between testings. Rawson and Dunlosky (2011, 2012) have shown in various experiments that testing each content three times, with feedback, is better than testing it once in terms of later recall. However, just like McDermott et al. (2014), Rawson and Dunlosky (2011, 2012) did not test different time intervals between testing opportunities. The former paper used pre-test testing, after-class testing, or testing just before the final exam, while the latter study repeated the tests three times in a row (massive testing), or three times in a spaced manner. In short, the literature does not have an answer on the ideal number of repetitions of retrieval practice attempts regarding a given content nor on the ideal time interval between these attempts.

If the student chooses to practice retrieval of a piece of information more than once, there are indications that there must be a little forgetfulness before each retrieval attempt, as this seems to be important to benefit the consolidation of memories (Roediger & Karpicke, 2011). This is in accordance with the theory of desirable difficulties, i.e., that some difficulty in recalling information improves long-term learning (Bjork & Bjork, 2011). Stated differently, there is less forgetfulness of content between repeated testing at short intervals so this leads to less effort in trying to recall something. Differently, more forgetfulness ensues if the interval between tests increases, leading to greater effort in trying to remember, which enhances learning. It is important to observe, nonetheless, that the literature is not clear about how much "forgetfulness" or "difficulty" is ideal, nor does it take into account that this can vary between different types of content.

As we have seen, the ideal number of retrieval practice attempts indicated by different authors is not consensual. In this sense, in order not to take up the short class time available to teach content, it is suggested that teachers encourage students to try to remember each study content at least once (Bangert-Drowns et al., 1991) in the classroom, because the greatest advantage of testing occurs with one retrieval practice attempt (Rowland, 2014). If students can practice retrieving specific pieces of information more often, that is even better. If the teacher chooses to repeat the retrieving practice sessions, he/she should do so in a distributed way over time, which is better than immediately consecutive retrieving practices, or massive repetition (Karpicke, 2017). The ideal time interval between exposure to content and trying to recall it, however, remains unknown. Therefore, until this is determined, retrieval practice must be carried out in a way that best suits each teacher.

When teachers choose to repeat testing opportunities, it should be considered that there are several possible schemes to be followed: there is the massive scheme (repeat several times consecutively over a short period) and spaced scheme (distributed over time). The latter can be subdivided into two types (see Karpicke & Roediger, 2010; Roediger & Karpicke, 2011): i) fixed (fixed time intervals between tests); ii) expanded (first test soon after the presentation of information, followed by testing over increasing time intervals); or iii) contraction (long intervals that become shorter over time). Another way to distribute tests over time is to adjust the intervals according to test difficulty so that tests on more challenging are repeated over shorter intervals, while easier tests are repeated at longer intervals (see Mubarak & Smith, 2008; Pham, Chen, Nguyen, & Hwang, 2016). However, it is not clear which of these schemes is the best, as there is still much controversy on this issue (e.g. Kang, Lindsey, Mozer, & Pashler, 2014; Karpicke & Bauernschmidt, 2011). This is not surprising, given that the study materials, number of tests, and intervals among test repetitions vary

considerably from publication to publication.

What is evident from prior studies is that massive testing schemes are disadvantageous for long-term retention compared to distributed retrieval practice over time (Karpicke & Roediger, 2007; Roediger & Karpicke, 2011), possibly because remembering information in repeated tests over a short period of time is easier: information is "fresh" in the student's memory (little difficulty) so that less recall effort ends up eliciting lower long-term retrieving practice effects (Cull, 2000). This corroborates the theory of desirable difficulties (Bjork & Bjork, 2011), mentioned above.

Both massive and spaced schemes, however, have a similar effect if the assessment of retention occurs as soon as the next day, for example (see Roediger & Pyc, 2012a). The realization of this fact by students sometimes leads them to postpone studying until one day before the examination. However, what they do not realize is that if they study like this they will have more difficulty remembering the content if they are questioned about the material later in time (Son, 2004). That is, trying to recall information repeatedly in a massively format, over a short time period, leads to less long-term learning than practicing retrieval in a spaced-out manner over days or weeks. Therefore, if possible, teachers should provide exercises on different content throughout the course, preferably mixing tests or activities about the content they are discussing in class with tests on topics that were studies before (cumulative tests).

In a nutshell, it is unknown what is the best scheme for the repetition of retrieval practice opportunities, except that doing so repeatedly over a short period of time, such as the evening before an exam, does not promote long-term learning. If the teacher is willing and able to administer repeated tests, it is suggested that he/she use the fixed scheme as it seems to present a slight advantage over other schemes (Schuetze, 2014). Moreover, it is easier for teachers to remember to repeat the tests always at the same time interval than to have to do this at increasing or decreasing intervals, as in the cases of expanded and contracted schemes.

FINAL CONSIDERATIONS

Given the need for education to be based on scientific evidence, the appropriation of knowledge of the factors that influence retrieval practice effects may have positive consequences in the implementation of this learning technique in the classroom. It should be kept in mind that the implementation of this way of teaching and studying does not involve additional costs (Roediger & Pyc, 2012b), an important issue considering the limitations in terms of financial resources in our country. To facilitate the application of the tips provided here, a summary table of what was discussed above can be found in Table 1. When using retrieval practice, teachers should always seek to provide feedback and offer opportunities to remember that involve effort in recalling the answers (neither too easy nor too difficult) and that vary as much as possible in terms of the types of question (such as multiple-choice, open-ended, or activities in which students ask each other questions on the learned material, etc.). Teachers should also strive to provide recall opportunities for each studied content at least once; if it is possible to do so more often, repeated testing should be distributed over time. Additionally, students should be encouraged to practice retrieval without becoming stressed out or being afraid of making mistakes.

Table 1. *Summary of factors that affect retrieval practice effects and evidence-based recommendations to educators.*

Factor	What it consists of	Recommendation for educators
Feedback	Tell students if answers were correct or not.	ALWAYS give students access to the correct answers to activities that involve the retrieval practice.
Test format and features	Example: multiple-choice, open-ended answer questions, fill-in-the-blank.	Try to provide retrieval practice opportunities for your students, preferably involving some cognitive effort (open-ended questions), either by doing exercises, or by asking them to come up with questions for classmates to answer.
Exam format/final exam	Exam formats may or may not be the same as those that provide opportunities for retrieving (e.g. multiple-choice, open-ended questions).	Ideally, exams and finals should not be in the same format as the retrieval practice attempts used to study the content. This allows the assessment of transfer of information to another context, in addition to creating another opportunity to practice retrieval.
Characteristics of test response modalities	Typed, spoken, written or covert answer modalities and open or closed book exams.	Response modality does not matter as long as students try to remember the content. In the cases of open book exams/test, do not inform students of this or, if you do, explain that if they study more they will find the answers in the support material more easily.
Number of repetitions of retrieval practice opportunities.	How often retrieval practice be repeated regarding each content.	Retrieval practice opportunities regarding each taught content should occur at least once. More repetitions increase memory, but returns are progressively lower and not always possible to do in real classroom contexts.
Intervals between retrieval practice opportunities.	The length of time between retrieval practice opportunities regarding the same content.	Preferably repeat retrieval practice opportunities regarding the same content in a distributed manner over time and not on the same day.

REFERENCES

Agarwal, P. K.; D'Antonio, L.; Roediger, H. L.; McDermott, K. B.; McDaniel, M. A. (2014). Classroom-based programs of retrieval practice reduce middle schooland high school students' test anxiety. *Journal of Applied Research in Memory and Cognition*, *3*(3), 131–139. https://doi.org/10.1016/j.jarmac.2014.07.002

Agarwal, P. K.; Karpicke, J. D.; Kang, S. H. K.; Roediger, H. L.; McDermott, K. B. (2008). Examining the Testing Effect with Open- and Closed-Book Tests. *Applied Cognitive Psychology*, *22*(7), 861–876. https://doi.org/10.1002/acp.1391

Agarwal, P. K.; Roediger, H. L. (2011). Expectancy of an open-book test decreases performance on a delayed closed-book test. *Memory*, *19*(8), 836–852. https://doi.org/10.1080/09658211.2011.613840

Agarwal, P. K.; Roediger, H. L.; McDaniel, M. A.; McDermott, K. B. (2013). *How to use retrieval practice to improve learning*. Retrieved from Washington University in St. Louis website: www.retrievalpractice.org

Bangert-Drowns, R. L.; Kulik, J. A.; Kulik, C.-L. C. (1991). Effects of frequent classroom testing. *The Journal of Educational Research*, *85*(2), 89–99. https://doi.org/10.2307/27540459

Bjork, E. L.; Bjork, R. A. (2011). Making Things Hard on Yourself, But in a Good Way: Creating Desirable Difficulties to Enhance Learning. In Gernsbacher, M. A.; Pew, R. W.; Hough, L. M.; Pomerantz, J. R. (Eds.), *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society* (pp. 55–64). https://doi.org/10.1017/CBO9781107415324.004

Bruer, J. T. (1997). Education and the brain: A bridge too far. *Educational Researcher*, *26*(8), 4–16.

Butler, A. C.; Karpicke, J. D.; Roediger, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology. Applied*, *13*(4), 273–281. https://doi.org/10.1037/1076-898X.13.4.273

Butler, A. C.; Karpicke, J. D.; Roediger, H. L. (2008). Correcting a Metacognitive Error: Feedback Increases Retention of Low-Confidence Correct Responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 918–928. https://doi.org/10.1037/0278-7393.34.4.918

Cantor, A. D.; Eslick, A. N.; Marsh, E. J.; Bjork, R. A.; Bjork, E. L. (2014). Multiple-choice tests stabilize access to marginal

knowledge. *Memory & Cognition*, *43*(2), 193–205. https://doi.org/10.3758/s13421-014-0462-6

Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology*, *14*(3), 215–235. https://doi.org/10.1002/(sici)1099-0720(200005/06)14:3<215::aid-acp640>3.0.co;2-1

de Jonge, M.; Tabbers, H. K.; Rikers, R. M. J. P. (2015). The Effect of Testing on the Retention of Coherent and Incoherent Text Material. *Educational Psychology Review*, *27*(2), 305–315. https://doi.org/10.1007/s10648-015-9300-z

Ekuni, R.; Pompeia, S. (2015). Prova é perda de tempo? In R. Ekuni, L. Zeggio, & O. F. A. Bueno (Eds.), *Caçadores de Neuromitos: o que você sabe sobre seu cérebro é verdade?* (pp. 196–202). São Paulo: Memnon.

Hays, M. J.; Kornell, N.; Bjork, R. A. (2010). The costs and benefits of providing feedback during learning. *Psychonomic Bulletin & Review*, *17*(6), 797–801. https://doi.org/10.3758/PBR.17.6.797

Hays, M. J.; Kornell, N.; Bjork, R. A. (2012). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 290–296. https://doi.org/10.1037/a0028468

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

Kang, S. H. K., Lindsey, R. V, Mozer, M. C., & Pashler, H. (2014). Retrieval practice over the long term: Should spacing be expanding or equal-interval? *Psychonomic Bulletin & Review*, *21*(6), 1544–1550. https://doi.org/10.3758/s13423-014-0636-z

Karpicke, J. D. (2017). Retrieval-Based Learning: A Decade of Progress. In Byrne, J. H. (Ed.), *Learning and Memory: A Comprehensive Reference* (Third Edit, pp. 487–514). https://doi.org/10.1016/B978-0-12-805159-7.02023-4

Karpicke, J. D., & Aue, W. R. (2015). The Testing Effect Is Alive and Well with Complex Materials. *Educational Psychology Review*, *27*(2), 317–326. https://doi.org/10.1007/s10648-015-9309-3

Karpicke, J. D., & Bauernschmidt, A. (2011). Spaced retrieval: Absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*(5), 1250–1257. https://doi.org/10.1037/a0023436

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*, *17*(4), 471–479. https://doi.org/10.1080/09658210802647009

Karpicke, J. D.; Roediger, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, *57*(2), 151–162. https://doi.org/10.1016/j.jml.2006.09.004

Karpicke, J. D.; Roediger, H. L. (2010). Is expanding retrieval a superior method for learning text materials? *Memory & Cognition*, *38*(1), 116–124. https://doi.org/10.3758/MC.38.1.116

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

Little, J. L.; Bjork, E. L.; Bjork, R. A.; Angello, G. (2012). Multiple-Choice Tests Exonerated, at Least of Some Charges: Fostering Test-Induced Learning and Avoiding Test-Induced Forgetting. *Psychological Science*, *23*(11), 1337–1344. https://doi.org/10.1177/0956797612443370

Marsh, E. J.; Agarwal, P. K.; Roediger, H. L. (2009). Memorial consequences of answering SAT II questions. *Journal of Experimental Psychology. Applied*, *15*(1), 1–11. https://doi.org/10.1037/a0014721

McDaniel, M. A.; Agarwal, P. K.; Huelser, B. J.; McDermott, K. B.; Roediger, H. L. (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.; Anderson, J. L.; Derbish, M. H.; Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19*(4–5), 494–513. https://doi.org/10.1080/09541440701326154

McDaniel, M. A.; Roediger, H. L.; McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, *14*(2), 200–206. https://doi.org/10.3758/BF03194052

McDermott, K. B.; Agarwal, P. K.; D'Antonio, L.; Roediger, H. L.; 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

Mubarak, R.; Smith, D. C. (2008). Spacing Effect and Mnemonic Strategies: a Theory-Based Approach To E-Learning. *E-Learning*, 269–272.

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

Pham, X.-L.; Chen, G.-D.; Nguyen, T.-H.; Hwang, W.-Y. (2016). Card-based design combined with spaced repetition: A new interface for displaying learning elements and improving active recall. *Computers & Education*, *98*(July), 142–156. https://doi.org/10.1016/j.compedu.2016.03.014

Putnam, A. L.; Roediger, H. L. (2012). Does response mode affect amount recalled or the magnitude of the testing effect? *Memory & Cognition*, *41*(1), 36–48. https://doi.org/10.3758/s13421-012-0245-x

Rawson, K. A.; Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How

much is enough? *Journal of Experimental Psychology: General*, *140*(3), 283–302. https://doi.org/10.1037/a0023956

Rawson, K. A.; Dunlosky, J. (2012). When Is Practice Testing Most Effective for Improving the Durability and Efficiency of Student Learning? *Educational Psychology Review*, *24*(3), 419–435. https://doi.org/10.1007/s10648-012-9203-1

Roediger, H. L.; Agarwal, P. K.; McDaniel, M. A.; McDermott, K. B. (2011). Test-enhanced learning in the classroom: long-term improvements from quizzing. *Journal of Experimental Psychology: Applied*, *17*(4), 382–395. https://doi.org/10.1037/a0026252

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. (2006a). 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

Roediger, H. L.; Karpicke, J. D. (2006b). 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

Roediger, H. L.; Karpicke, J. D. (2011). Intricacies of spaced retrieval: A resolution. In Benjamin, A. S. (Ed.), *Successful Remembering and Successful Forgetting: Essays in Honor of Robert A. Bjork* (pp. 23–48). https://doi.org/10.4324/9780203842539

Roediger, H. L.; Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(5), 1155–1159. https://doi.org/10.1037/0278-7393.31.5.1155

Roediger, H. L.; Putnam, A. L.; Smith, M. A. (2011). Ten Benefits of Testing and Their Applications to Educational Practice. In Mestre, J.; Ross, B. (Ed.), *Psychology of learning and motivation: Cognition in education* (Vol. 55, pp. 1–36). https://doi.org/10.1016/B978-0-12-387691-1.00001-6

Roediger, H. L.; Pyc, M. A. (2012a). Applying cognitive psychology to education: Complexities and prospects. *Journal of Applied Research in Memory and Cognition*, *1*(4), 263–265.

Roediger, H. L.; & Pyc, M. A. (2012b). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition*, *1*(4), 242–248.

Rowland, C. A. (2014). The Effect of Testing Versus Restudy on Retention: A Meta-Analytic Review of the Testing Effect. *Psychological Bulletin*, *140*(6), 1–32. https://doi.org/10.1037/a0037559

Schuetze, U. (2014). Spacing techniques in second language vocabulary acquisition: Short-term gains vs. long-term memory. *Language Teaching Research*, *19*(1), 1–15. https://doi.org/10.1177/1362168814541726

SCImago. (2019). SIR Ranking of Brazil. SJR — *SCImago Journal & Country Rank*. https://www.scimagojr.com/countrysearch.php?country=BR

Slavin, R. E. (2020). How evidence-based reform will transform research and practice in education. *Educational Psychologist*, *55*(1), 21–31. https://doi.org/10.1080/00461520.2019.1611432

Smith, M. A.; Roediger, H. L.; Karpicke, J. D. (2013). Covert retrieval practice benefits retention as much as overt retrieval practice. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(6), 1712–1725. https://doi.org/10.1037/a0033569

Smith, T. A.; Kimball, D. R. (2010). Learning from Feedback: Spacing and the Delay-Retention Effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 80–95. https://doi.org/10.1037/a0017407

Son, L. K. (2004). Spacing one's study: evidence for a metacognitive control strategy. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(3), 601–604. https://doi.org/10.1037/0278-7393.30.3.601

Stern, E. (2005). Pedagogy meets neuroscience. *Science*, *310*, 745. https://doi.org/10.1126/science.1121139

UNESCO. (2014). *Education index*. http://hdr.undp.org/en/content/education-index.

van den Broek, G.; Takashima, A.; Wiklund-Hörnqvist, C.; Karlsson Wirebring, L.; Segers, E.; Verhoeven, L.; Nyberg, L. (2016). Neurocognitive mechanisms of the "testing effect": A review. *Trends in Neuroscience and Education*, *5*(2), 52–66. https://doi.org/10.1016/j.tine.2016.05.001

Weinstein, Y.; McDermott, K. B.; Roediger, H. L. (2010). A comparison of study strategies for passages: rereading, answering questions, and generating questions. *Journal of Experimental Psychology. Applied*, *16*(3), 308–316. https://doi.org/10.1037/a0020992

Weinstein, Y.; Roediger, H. L. (2010). Retrospective bias in test performance: Providing easy items at the beginning of a test makes students believe they did better on it. *Memory & Cognition*, *38*(3), 366–376. https://doi.org/10.3758/MC.38.3.366

Weinstein, Y.; Roediger, H. L. (2012). The effect of question order on evaluations of test performance: how does the bias evolve? *Memory & Cognition*, *40*(5), 727–735. https://doi.org/10.3758/s13421-012-0187-3