

Prática de lembrar em ambiente escolar: revisão narrativa sobre formatos de testes

Práctica de recuperación en escuelas: revisión narrativa sobre formatos de las pruebas

Retrieval practice at school: a narrative review of test formats

Anne Caroline Borba da Silvaⁱ

Universidade Estadual do Norte do Paraná. Jacarezinho, PR, Brasil

annecarolsap@gmail.com

<https://orcid.org/0000-0002-0689-1196>

Lisandra Tamires Mendonçaⁱⁱ

Universidade Estadual do Norte do Paraná. Jacarezinho, PR, Brasil

lisandratamiresm@gmail.com

<https://orcid.org/0000-0002-7880-5955>

José Cláudio de Sene Miguelⁱⁱⁱ

Universidade Estadual do Norte do Paraná. Jacarezinho, PR, Brasil

joseclaudiosm@gmail.com

<https://orcid.org/0000-0001-6097-2950>

Roberta Ekuni^{iv}

Universidade Estadual do Norte do Paraná. Jacarezinho, PR, Brasil

robertaekuni@uenp.edu.br

<https://orcid.org/0000-0003-1490-0184>

Os autores contribuíram igualmente na elaboração do manuscrito.

Recebido em: 21/04/2022

Aceito em: 28/07/2022

Publicado em: 16/08/2022

Linhas Críticas | Periódico científico da Faculdade

de Educação da Universidade de Brasília, Brasil

ISSN: 1516-4896 | e-ISSN: 1981-0431

Volume 28, 2022 (jan-dez).

<http://periodicos.unb.br/index.php/linhascriticas>

Referência completa (APA):

Silva, A. C. B. da, Mendonça, L. T., Miguel, J. C. de

S., & Ekuni, R. (2022). Prática de lembrar em

ambiente escolar: revisão narrativa sobre

formatos de testes. *Linhas Críticas*, 28, e43014.

<https://doi.org/10.26512/lc28202243014>

Link alternativo:

<https://periodicos.unb.br/index.php/linhascriticas/article/view/43014>

Licença Creative Commons  CC BY 4.0.

Resumo: A prática de lembrar, que consiste em tentar recordar conteúdo já visto, é tida como uma das estratégias mais eficazes para gerar aprendizado de longa duração. A fim de promover seu uso eficaz, a presente revisão narrativa tem por foco sugerir diretrizes a respeito da aplicação de diferentes formatos de testes, com frequência empregados para praticar lembrar, e apontar vantagens e desvantagens de cada um. Como resultado, os estudos revisados envolveram alunos desde a educação infantil até a graduação. No geral, a prática de lembrar pode beneficiar a aprendizagem de diferentes conteúdos contanto que os formatos de teste se adaptem à idade do aprendiz, independentemente dos materiais usados.

Palavras-chave: Educação. Formato de teste. Aprendizagem.

Resumen: La práctica de recuperación, es decir, intentar recordar contenido ya visto, es considerada una de las estrategias más efectivas para generar aprendizaje de larga duración. Para promover su uso eficaz, esta revisión narrativa se enfoca en sugerir pautas sobre la aplicación de diferentes formatos de pruebas, a menudo usados para practicar recordar, y señalar sus ventajas y desventajas. Los estudios revisados involucraron a alumnos desde el jardín de infantes hasta la graduación. En general, la práctica de recuperación puede beneficiar el aprendizaje de diferentes contenidos siempre que los formatos de prueba se adapten a la edad del alumno, independientemente de los materiales utilizados.

Palabras clave: Educación. Formato de prueba. Aprendizaje.

Abstract: Retrieval practice, which consists of trying to remember content already seen, is considered as one of the most effective strategies to generate long-term learning. In order to promote its effective use, this narrative review focuses on suggesting guidelines for the application of different test formats, often used to practice retrieval, and pointing out advantages and disadvantages of each. As a result, the reviewed research encompassed students from kindergarten through undergraduation. Overall, retrieval practice can benefit the learning of different content, providing test formats are suitable to the learner's age, regardless of the materials used.

Keywords: Education. Test format. Learning.

ⁱ Mestranda em Educação Básica pela Universidade Estadual do Norte do Paraná.

ⁱⁱ Mestra em Educação Básica pela Universidade Estadual do Norte do Paraná (2021).

ⁱⁱⁱ Mestrando em Educação Básica pela Universidade Estadual do Norte do Paraná.

^{iv} Professora Adjunta no Programa de Pós-graduação em Educação da Universidade Estadual do Norte do Paraná.

Introdução

Um dos objetivos da educação é que a aprendizagem persista ao longo do tempo. Todavia, é comum haver queixas de estudantes a respeito de não conseguirem se lembrar do conteúdo estudado. Um dos motivos é que eles não preferem utilizar as melhores estratégias para aprender (Karpicke et al., 2009; Dunlosky et al., 2013; Karpicke et al., 2014; Ekuni et al., 2020). Diante disso, diversas pesquisas na área da Psicologia Cognitiva, tanto em laboratórios quanto em contextos educacionais, apontam estratégias que promovem aprendizagem de longa-duração, como a prática de lembrar (Dunlosky et al., 2013).

A prática de lembrar (*retrieval practice*) é uma estratégia de aprendizagem que visa a tentar lembrar dos conteúdos vistos anteriormente, seja por meio de testes (múltipla escolha, complete, resposta curta etc.), seja por meio de exercícios que estimulem lembrar (Roediger & Karpicke, 2006a). Durante a codificação da informação (que ocorre, por exemplo, ao assistir a uma aula, ao realizar a leitura de um texto etc.), a informação é colocada dentro de nossa cabeça. Ao tentar lembrar, porém, buscamos essa informação, tentando “puxá-la” para fora da cabeça. Nessa tentativa, o cérebro elabora e relaciona rotas (caminhos) disponíveis para identificar informações a que tivemos acesso anteriormente (Bjork, 1975) e ativa conteúdos semanticamente relacionados (Carpenter, 2011). Isso quer dizer que, quando encontramos o conteúdo que procuramos em nossa mente, a informação é reconsolidada, seu acesso é facilitado, e o traço de memória é fortalecido, tornando-o mais duradouro (Van den Broek et al., 2016).

Historicamente, experimentos que testam os efeitos da prática de lembrar datam de mais de um século. Os artigos revisados convergem à pesquisa conduzida por Abbott (1909), indicado como o marco inicial dos estudos na área (Yang et al., 2021). A descoberta de Abbott, depois replicada por inúmeros pesquisadores, é a de que testar o conhecimento aprendido altera sua retenção na memória (Abbott, 1909; Roediger & Karpicke, 2011). A ebulição de pesquisas sobre prática de lembrar, contudo, sobreviria só nas décadas de 1960 e 1970, a partir da publicação de outros artigos relevantes a respeito de temas pertinentes à Psicologia Cognitiva da Memória (Roediger & Karpicke, 2011). Apesar de um século de pesquisas na área, Agarwal e Bain (2019) esclarecem que as “modas pedagógicas” (*fads*) são mostradas em cursos de formação de professores e acabam por preterir as estratégias de ensino baseadas em evidências. Talvez isso explique o motivo pelo qual a prática de lembrar não é amplamente vista como estratégia de ensino no cotidiano escolar.

Praticar lembrar, entre seus benefícios, melhora a metacognição (McDermott, 2021), diminui a ansiedade em testes posteriores (Agarwal et al., 2014), fornece *feedback* sobre o que se sabe e o que não se sabe, abrindo espaço, desse modo, a estudos futuros (Roediger & Karpicke, 2006b). Outra vantagem é que não envolve investimento financeiro adicional (Roediger & Pyc, 2012) e pode-se adaptar a diversos materiais educacionais e métodos de ensino (Agarwal et al., 2018).

Como dito, uma das formas de praticar lembrar pode dar-se por meio de testes, pois, quando o aluno vê uma determinada questão, tem de tentar extrair a resposta de sua cabeça. Todavia, na

prática educacional, geralmente os testes são usados como forma de avaliação, e não como estratégia de aprendizagem (Yang et al., 2021). A implementação da prática de lembrar na educação é importante no sentido de promover uma aprendizagem duradoura, de qualidade, e de estimular os alunos a que se dediquem aos seus próprios processos cognitivos por meio de estratégias de estudo efetivas (Ekuni et al., 2020).

A educação precisa de métodos comprovados cientificamente e que possibilitem efeitos concretos no ambiente escolar (Slavin, 2020). Portanto, o uso de estratégias de aprendizagem mais efetivas pode beneficiar os estudantes. Diante o exposto, o presente estudo teve como objetivo realizar uma revisão narrativa, estilo de revisão que não exige que se siga um protocolo sistemático, de modo que não é necessário informar a metodologia de busca das referências (Rother, 2007). Tal revisão pode ser feita pelo(s) o(s) próprio(s) autor(es), sem que seja necessário explicitar os critérios das buscas e seleção das fontes (Collins & Fauser, 2005). Todavia, deve-se garantir eloquência acadêmica a partir da análise crítica do(s) autor(es) a respeito da literatura publicada (Rother, 2007). Igualmente, deve-se cuidar para que as fontes selecionadas sejam baseadas em estudos científicos comprovadamente eficazes (Slavin, 2020). Seguindo essas diretrizes, é possível estabelecer uma relação entre as produções sobre temas previamente definidos, consolidar conceitos e fornecer orientações práticas (Elias et al., 2012). Assim, nas revisões narrativas, as produções bibliográficas são analisadas de tal maneira que resultem num estado da arte sobre o tema em que o pesquisador pretende se aprofundar (Elias et al., 2012).

Para analisar o uso de diferentes formatos de testes como estratégia de aprendizagem, conforme explicitado acima, utilizamos como método a revisão narrativa. Como critérios de inclusão utilizamos artigos que focaram testes como estratégia de aprendizagem, e não como instrumento avaliativo. Com efeito, o objetivo do uso dessas estratégias não é avaliar o aluno, mas, sim, incentivá-lo a se esforçar para tentar lembrar do conteúdo aprendido. Nesse sentido, a literatura conhece esse fenômeno como *testing-effect* (efeito da testagem), *test-enhanced learning* (aprendizagem melhorada por teste), *retrieval-based learning* (aprendizagem baseada na lembrança), *retrieval practice* (prática de lembrar), termos utilizados nas buscas realizadas em bases de dados de forma não sistemática. Por fim, tratamos de subsídios para que os professores e alunos implementem a prática de lembrar no processo de ensino e aprendizagem.

Verdadeiro ou falso (*true-false*)

As questões de verdadeiro ou falso (V/F) são questões de seleção de resposta, ou seja, o respondente deve reconhecer a resposta a ser selecionada, decidindo se uma afirmação é falsa ou verdadeira (Santrock, 2009). Por exemplo: *O gêiser Castle é o gêiser mais velho do parque Yellowstone: (V) / (F)*.

Do ponto de vista de suas vantagens, questões desse tipo são comumente usadas em sala de aula e pedagogicamente práticas de serem administradas, pois são objetivas na mensuração de

resultados (pontuação com base em acertos) e demandam menos tempo de preparo (Uner et al., 2022). Ademais, esse tipo de teste permite cobrir um grande número de questões num período comum de aplicação de testes (Santrock, 2009). Se houver situações em que seja difícil criar vários distratores, ou seja, alternativas plausíveis para as questões de múltipla escolha, esse tipo de teste também é útil (Uner et al., 2022).

Testes de V/F resultam em benefícios também com crianças. Por exemplo, uma pesquisa que envolveu estudantes (idade média de onze anos) em ambiente de sala de aula, apontou efeitos positivos na performance cognitiva (correção das respostas) e metacognitiva (nível de confiança nas respostas) a partir de experimentos que envolviam questões de V/F sobre conteúdos de história, política e geografia (Barenberg & Dutke, 2019). Conforme indicam experimentos realizados com quase quinhentos participantes universitários, empregando-se conteúdos de biologia em ambientes de aprendizagem virtual, as questões de V/F também podem ser benéficas em cenários de aprendizagem remota (*e-learning*) ou mesclados com o ensino presencial, sobretudo quando há a presença de *feedback* (Enders et al., 2021).

Em sentido contrário, a eficácia pedagógica dos testes de V/F pode não ser tão poderosa em comparação a outros formatos de teste, conforme apontam experimentos conduzidos com estudantes universitários, em ambiente de laboratório, em que se usaram textos (Brabec et al., 2021). Assim, entre as possíveis desvantagens da aplicação desse tipo de teste, figuram os chutes (isto é, tentativas de acerto) e a sugestão negativa. Neste último caso, o estudante, ao ler o distrator (informação falsa) pode aprendê-lo e reconhecê-lo futuramente em outros testes, tomando-o por verdadeiro (Uner et al., 2022). Nessa esteira, em experimentos conduzidos com universitários, Brabec et al. (2021) encontraram evidências de sugestão negativa quando os participantes tinham de escolher a alternativa verdadeira em comparação a não serem testados com perguntas de V/F (condição-controle). De acordo com os autores, esse efeito negativo pode ser minorado com a presença de *feedback* (Brabec et al., 2021).

Para tentar minimizar esse senão, bem como para evitar chutes, um experimento conduzido com estudantes de Psicologia, solicitou que os participantes fornecessem uma justificativa para as respostas selecionadas (Schaap et al., 2014). Como se vê, trata-se de um teste de V/F ligeiramente modificado. No entanto, essa estratégia não gerou benefício significativo em comparação aos participantes que não haviam de justificar suas escolhas (Schaap et al., 2014). Em atenção a isso, outros experimentos modificados demandaram de participantes universitários não que justificassem as escolhas, mas que corrigissem as alternativas falsas (Uner et al., 2022).

Outras modificações experimentadas envolveram a inserção de frases parentéticas (*competitive clauses*) nas afirmativas, estratégia que talvez possa fazer perdurar a informação testada por mais tempo na memória (Brabec et al., 2021). Por exemplo, em vez de a afirmativa figurar sozinha, há inserida nela outra afirmação, entre parênteses: o gêiser *Castle* (e não o *Stemboat*) é o gêiser mais alto (exemplo retirado do artigo supramencionado).

É relevante fazer constar que quando os estudantes são levados a corrigir as alternativas que julgam ser falsas, a apresentação de *feedback* melhora a retenção de itens corretos num teste posterior de respostas curtas (*short-answer*) em comparação a um procedimento regular de V/F (Uner et al., 2022). Há três tipos de *feedback*: o primeiro, do tipo “certo ou errado”; o segundo, chamado *feedback* corretivo, que provê o estudante com o teor da resposta correta; e o terceiro, chamado *feedback* elaborado, que explica por que razão certa resposta é verdadeira ou falsa (Enders et al., 2021).

O *feedback* corretivo pode melhorar a retenção de itens testados em questões de verdadeiro ou falso. Igualmente, simples modificações (tais como inserir frase parentética e solicitar correção da alternativa errada ou justificativa da escolha) na forma de administrar o teste parecem ser mais eficazes e promissoras para implementação em sala de aula (Uner et al., 2022).

No que diz respeito a sugestões para elaboração de questões de V/F, a recomendação aos professores é que fiquem atentos a alguns princípios: manter apenas uma ideia principal em cada alternativa, e não inserir várias ao mesmo tempo; fazer afirmações curtas e de vocabulário compreensível; evitar termos absolutos (*sempre, nunca, ninguém* etc.), modalizadores circunstanciais (*pode, talvez, às vezes* etc.) e duplas negativas (Sanrock, 2009).

Múltipla escolha (*multiple choice*)

As questões de múltipla escolha (ME) também envolvem seleção de resposta. No tocante à forma, esse tipo de questão é composto por duas partes: a base (enunciado) e seu conjunto de respostas possíveis, das quais apenas uma é a correta, e as outras são *distratores* (Sanrock, 2009). Testes de ME são frequentemente usados por educadores porque, assim como V/F, permitem pontuação mais fácil e são percebidos como mais objetivos (Butler et al., 2006). Na verdade, trata-se do tipo de teste objetivo mais comumente usado (Marsh et al., 2007). Ademais, questionários de ME podem ser empregados ao longo do período letivo para beneficiar o aprendizado e o rendimento dos estudantes nos testes finais, sejam estes de ME (composto por questões de reconhecimento) ou de respostas curtas (composto por questões de elaboração) (McDermott et al., 2014).

Questões de ME devem ser redigidas de modo que os distratores sejam plausíveis (Little et al., 2012). Porém, assim como já mencionado no excerto sobre questões do tipo V/F, há o inconveniente efeito da sugestão negativa. Os distratores (respostas incorretas) podem ser aprendidos e reproduzidos em teste posterior, de modo que o teste de ME pode ser contraproducente se houver muitos deles, pois sua interferência pode levar à assimilação de conhecimento errôneo (e.g. Marsh et al., 2012a). Com efeito, um experimento conduzido com universitários aponta que quanto maior o número de alternativas oferecidas no teste de ME, tanto menor é o benefício do teste (Roediger & Marsh, 2005). Além da interferência causada por muitas alternativas distratoras, o estudo insuficiente do conteúdo anterior ao teste contribui para a aprendizagem de informações errôneas (Butler & Roediger, 2008).

Para evitar a sugestão negativa e potencializar os benefícios dos testes de ME é importante fornecer *feedback* (Marsh et al., 2012a). O *feedback*, particularmente o de correção (em que a resposta correta é dada), é útil para corrigir informações errôneas que os alunos tenham aprendido dos distratores, bem como para manter a retenção em longo prazo das respostas corretas (Butler, 2018).

Testes de ME foram empregados em várias pesquisas cuja idade, *setting* e materiais empregados eram diferentes. Experimentos conduzidos com crianças (idade média: oito anos) demonstram que elas se beneficiam de testes de ME sobre conhecimentos gerais, associados a *feedback* (Marsh et al., 2012b). Outros experimentos com adolescentes (idade média: treze anos) apontam que, mesmo que sejam breves, questionários de ciências realizados em sala de aula acompanhados de *feedback* aumentam a performance de estudantes em avaliações somativas posteriores, sejam elas de ME ou de resposta curta (McDermott et al., 2014).

Quanto a sugestões de como elaborar questões de ME, podem-se elencar algumas de cunho geral e outras específicas desse tipo de teste (Santrock, 2009; Butler, 2018). No geral, questões de ME não podem conter impropriedades gramaticais, devem ser compreensíveis, são melhores se redigidas como frases interrogativas e não devem conter alternativas capciosas. Especificamente, devem ter respostas de extensão similar e alternar a posição das alternativas corretas entre as questões. Convém repisar que quanto maior o número de alternativas, menor tende a ser a taxa de acerto (Butler et al., 2006). De mais a mais, testes de ME devem ser adequadamente desafiadores, pois se muito fáceis ou muito difíceis são inúteis tanto para avaliação quanto para promoção de aprendizagem (Butler, 2018). Uma sugestão pertinente é a opção “não sei”, de modo que a evitar chutes (Marsh et al., 2007). Isso porque o objetivo de aplicar testes é a aprendizagem, e não o fornecimento de nota.

Como exposto anteriormente, entre suas principais desvantagens estão a elaboração das alternativas (tanto no tocante ao número, quanto ao formato, quanto ainda ao grau de dificuldade) e a possibilidade de haver sugestão negativa (Butler, 2018). Por outro lado, esse formato de teste apresenta a vantagem de ser facilmente conferível. Desse modo, diminui o tempo de correção e aumenta sua objetividade (Butler et al., 2006; Marsh et al., 2012a). Com efeito, pode ser aplicado ao longo de todo o período letivo para impulsionar a aprendizagem (McDermott et al., 2014). Por fim, o conhecimento marginal — isto é, aquele que, mesmo armazenado na memória, não é acessível em dado momento — pode ser facilmente reativado por meio de testes de ME (Cantor et al., 2015).

Questões de ME têm, portanto, implicação pedagógica importante, seja em questionários, seja em testes finais, pois são práticas e econômicas, mas exigem cuidado em sua elaboração. A recomendação aos professores, pois, é cuidar do número de alternativas e procurar elaborar questões claras e alternativas plausíveis, com nível de dificuldade adequado aos alunos.

Resposta curta (*short answer*)

O formato de teste de resposta curta (RC) é muito utilizado em sala de aula. Como se pode depreender de seu nome, equivale a responder perguntas com uma breve resposta. Trata-se de um teste de produção, ou seja, impõe ao aluno que produza uma resposta (Larsen et al., 2008). Esse tipo de teste é semelhante a testes com pistas. Nesse caso, porém, a questão é uma pista que direciona o respondente ao conteúdo de que se deve lembrar (Moreira et al., 2019). Por exemplo, após estudarem definições de termo-chave, solicitou-se a alunos do 5º ano que digitassem as respostas [e.g. O que é som? ____ (*forma de energia que você pode ouvir e que viaja através da matéria como ondas*)] (Lipko-Speeda et al., 2014).

Esse formato de teste, assim como o teste de recordação livre, exige maior esforço para lembrar e é mais difícil do que o formato de teste de múltipla escolha, que envolve reconhecimento (Rowland, 2014). Efetivamente, pesquisas apontam que testes de RC são mais eficientes na retenção de conteúdo do que os de múltipla escolha (e.g. Kang et al., 2007; Stenlund et al., 2016). Entretanto, há estudos que divergem nesse ponto (ver Little et al., 2012). Em relação a essa discordância, é importante considerar a faixa etária dos alunos e a maneira segundo a qual os testes são formulados. Os testes devem ser elaborados de modo que os alunos tentem lembrar de fato em vez de somente identificar as respostas (Little et al., 2012).

Diversos estudos com diferentes níveis educacionais foram realizados com o formato de teste de RC, desde o Ensino Fundamental (e.g. Goossens et al., 2016), passando pelo Ensino Médio (Dirkx et al., 2014), até a Graduação (e.g. Endres et al., 2020). Essas pesquisas utilizaram como materiais: capítulos de livros (e.g. Carpenter et al., 2009), listas de palavras (Goossens et al., 2016), pares-chave de definição de conceitos (e.g. Lipko-Speeda et al., 2014), textos expositivos (e.g. Dirkx et al., 2014), conceitos-chave de tópicos estudados (Wiklund-Hörnqvist et al., 2014), aulas expositivas (Foss & Pirozzolo, 2017) e palestras (Lyle & Crawford, 2011). Com exceção do estudo de Goossens et al. (2016), em todos os supracitados houve benefícios do teste de resposta curta para a aprendizagem.

Lipko-Speeda et al. (2014) analisaram o efeito de realizar teste de RC com e sem *feedback*, empregando a releitura como condição-controle. O público-alvo foi composto por crianças de 5º ano, e as perguntas consistiam em definições de conceitos-chave de conteúdos de Ciências e Geografia. Os efeitos positivos foram vistos apenas na condição de teste com *feedback*. Em seu estudo realizado com crianças, Goossens et al. (2016) constataram que esse formato de teste sem *feedback* não foi melhor do que a cópia do material estudado. De forma semelhante, uma pesquisa conduzida com alunos do Ensino Médio revelou que o teste de RC com *feedback*, comparado à releitura e administrado após estudo de um texto sobre probabilidade, aumentou a retenção de conteúdo do texto e melhorou a aplicação dos princípios abrangidos no texto (Dirkx et al., 2014).

As pesquisas com alunos de graduação apontaram que, comparados à releitura, os testes de RC foram benéficos para a aprendizagem e retenção a longo prazo (Larsen et al., 2009; Wiklund-Hörnqvist et al., 2014; Greving & Richter, 2018). Um estudo de Carpenter et al. (2016) mostrou que

os testes de RC geraram melhores resultados no que diz respeito à recordação de definições de termos feita por alunos de alto desempenho. Todavia, para alunos de médio e baixo desempenho, copiar definições de termos foi melhor. Em todos esses estudos, foi oferecido *feedback* após o teste inicial de RC, com exceção do estudo de Greving e Richter (2018). Este último evidenciou que, mesmo sem *feedback*, o teste de resposta curta gerou efeitos positivos em conteúdo introdutório de Psicologia Cognitiva.

Testes de RC com *feedback* geram mais resultados na aprendizagem (Kang et al., 2007). Nesse sentido, por esse formato envolver mais esforço de recuperação, um dos cuidados mais importantes por parte dos professores é fornecer *feedback* com as respostas corretas. Assim sendo, os alunos podem corrigir os erros em vez de somente saber o que está certo ou errado.

As pesquisas citadas anteriormente mostraram que testes de RC com alunos mais velhos foram mais eficientes (e.g. Greving & Richter, 2018) do que com alunos mais novos (e.g. Lipko-Speeda et al., 2014). Recomenda-se, portanto, que, para alunos mais jovens, sejam fornecidas pistas (vide seção abaixo – recordação com pista) para facilitar a recordação. Outra saída é oferecer aos alunos mais oportunidades de aprendizagem e testes até que consigam integrar e reter o conteúdo (Lipko-Speeda et al., 2014).

Um dos benefícios de questões de RC é que elas favorecem a retenção de pontos específicos de um conteúdo estudado anteriormente. Por consequência, oportunizam a recordação de pontos mais difíceis ou pouco acessíveis. Como efeito indireto, possibilitam a calibração metacognitiva, isto é, a regulação da confiança dos alunos na certeza de suas respostas e na proporção de respostas dadas de que se lembrariam em uma semana. Desse modo, os julgamentos dos alunos em relação a sua aprendizagem são mais calibrados (precisos) no formato de RC em comparação ao formato de teste de recordação livre (Endres et al., 2020). Outra vantagem é que realizar revisão ao final de cada aula, por meio de questões de RC, incentiva o aluno a se engajar no estudo e aumenta seu desempenho em testes posteriores (Greving & Richter, 2018).

Uma das desvantagens dos questionários de RC são que, na prática, esse formato não é muito atrativo para os alunos. Além disso, os professores podem demorar até o dobro do tempo para corrigi-lo e aplicá-lo em sala de aula, pelo fato de o formato exigir respostas mais complexas (McDermott et al., 2014). Todavia, não é necessário corrigir individualmente, pois o *feedback* pode ser coletivo (Butler & Roediger, 2007), já que o objetivo é a aprendizagem, e não atribuição de nota em função do desempenho dos alunos. Outra desvantagem diz respeito ao fato de que sem o *feedback*, ou sem mais oportunidades de estudo, semelhante formato pode não ser tão eficaz para a retenção de informações.

Em suma, testes de RC são considerados importantes para a aprendizagem e recordação, haja vista que, por meio desse formato, os alunos podem lembrar fatos, definições de conceitos-chave e conteúdos específicos estudados. A recomendação aos professores é: sempre que possível, forneça o *feedback* a fim de que a integração do conteúdo seja mais eficaz.

Recordação livre (*free-recall*)

Testes que estimulam a recordação livre (RL) podem impulsionar de maneira significativa novas aprendizagens. Tal formato de teste visa a buscar, na mente, informações e conteúdos a que os estudantes tiveram acesso anteriormente, sem que lhes sejam fornecidas pistas para chegar à resposta correta (Brojde & Wise, 2008). Um exemplo é perguntar aos alunos o que eles lembram sobre o tema “sistema solar”.

Os testes de RL apresentaram resultados benéficos em diferentes níveis de ensino e faixas etárias, com crianças a partir de dois anos e meio de idade (Cornell et al., 1988), até jovens e adultos (Tulving, 1967). Entretanto, no teste final da pesquisa realizada por Aslan e Bäuml (2016), identificou-se que crianças mais jovens (idade média: seis anos) cometem mais erros ao utilizar a RL para praticar lembrar. Crianças mais velhas (idade média: oito anos), por sua vez, são mais beneficiadas com o uso de teste de RL (Aslan & Bäuml, 2016).

Quanto aos *settings* de pesquisas sobre a prática de lembrar, considerando apenas os testes em formato de RL, foram observados estudos realizados na casa dos participantes (Cornell et al., 1988), no laboratório (e.g. Lipowski et al., 2014) e na escola, em contexto de sala de aula (e.g. Jones et al., 2016).

Pesquisas mostram que, comparada à releitura dos mesmos conteúdos, a RL influencia e aprimora o processamento na recuperação de itens individuais (Tulving, 1967) e específicos (Lipowski et al., 2014). Ao comparar os efeitos da RL (prática de lembrar) com nenhum teste (Brojde & Wise, 2008; Roediger et al., 2011b) ou com estratégias de estudo comumente usadas, como a cópia (reescrita) (Jones et al., 2016; Rowley & McCrudden, 2020) e a releitura (Cornell et al., 1988; Aslan & Bäuml, 2016), observa-se que a prática de lembrar é uma das estratégias de aprendizagem mais eficazes e que proporciona aprendizagens mais duradouras (Roediger et al., 2011a). Assim, as vantagens de utilizar esse tipo de teste são que ele possibilita que a memória seja fortalecida (Cornell et al., 1988), promove a compreensão do conteúdo apresentado (Brojde & Wise, 2008), melhora a precisão ortográfica (Jones et al., 2016) e, conforme identificaram Rohrer et al. (2010), estimula a transferência de informações para novos contextos, beneficiando as aprendizagens de maneira robusta.

Nesse sentido, os testes de RL mostram um efeito direto do teste ao impulsionar os estudantes a lembrarem informações sem que sejam fornecidas pistas, estimulando, assim, uma dificuldade desejável. Isso beneficia especialmente as aprendizagens de crianças mais velhas. A recomendação aos professores é: forneça oportunidade para os alunos recordarem livremente o conteúdo (por exemplo, solicite que os alunos escrevam ou comentem o que lembram a respeito do conteúdo abordado na aula anterior, sem consultar material algum) para os estudantes a partir do Ensino Fundamental.

Teste de complete (fill-in-the-blank)

O teste de complete (TC) ou de preenchimento de lacunas é um tipo de teste que possibilita recordar uma ou determinadas palavras-chave (Hinze & Wiley, 2011). Por exemplo, no estudo de Jaeger et al. (2015), os alunos completaram frases após estudar um texto sobre o sol [e.g. *A palavra sol deriva da palavra latina _____ (solis)*].

Esse formato de teste é, numericamente, o menos explorado na literatura. Entretanto, os estudos existentes apontaram que essa tarefa possibilita recuperar itens de informações estudadas anteriormente. Consoante alguns estudos, o TC pode auxiliar a recuperação de palavras-chave de um texto enciclopédico com alunos de 3º ano (Jaeger et al., 2015), a aprendizagem de vocabulário de língua inglesa com alunos de 9º ano (Barenberg et al., 2021) e informações de itens sobre o desenvolvimento, estas apresentadas por meio de *PowerPoint* a graduandos de Psicologia (Vojdanoska et al., 2010).

Na pesquisa de Jaeger et al. (2015), os alunos que recordaram inicialmente com TC [e.g. *A camada superficial do sol é chamada de _____ (fotosfera)*] — tiveram um desempenho melhor no teste final de múltipla escolha após sete dias comparativamente aos alunos que releram as frases completas. Os autores argumentam, ainda, que a prática de lembrar por meio de TC pode beneficiar crianças que apresentam diferentes desempenhos de QI (quociente intelectual) e de habilidade de leitura.

Barenberg et al. (2021) realizaram experimentos com pares de palavras em alemão e inglês. De início, administraram o seguinte TC: uma pista (palavra alemã) e a palavra-alvo (palavra em inglês). Após uma semana, os alunos que passaram por esse tipo de teste apresentaram melhores resultados relativamente aos que passaram por releitura, seja quando realizaram teste final idêntico ao inicial, seja quando o fizeram em formato inverso (do idioma-alvo para o idioma-base).

O *feedback* esteve presente nas pesquisas de Barenberg et al. (2021) e Vojdanoska et al. (2010). Os resultados destes últimos revelaram que, fornecido o *feedback*, as vantagens do teste ampliaram-se em comparação ao teste sem *feedback* e sem atividade alguma.

Um dos pontos positivos de TC é ser simples de aplicar em sala de aula, fácil de corrigir e pouco demorado (Moreira et al., 2019). No entanto, há indícios de que esse formato pode não apresentar altos efeitos de prática de lembrar. Por exemplo, no experimento conduzido por de Jonge et al. (2015), os alunos de graduação estudaram um texto [coerente e não coerente (frases isoladas)] e realizaram um TC. Como condição-controle, foi utilizada a releitura das frases. Os autores concluíram que o TC foi mais benéfico para o formato de texto não coerente do que para o coerente. Assim, esse resultado pode ter-se decorrido do fato de esse tipo de tarefa não requerer integração e construção de conteúdo (Karpicke & Aue, 2015). Por outras palavras, esse tipo de tarefa não possibilita que o aluno processe e desenvolva significativamente mais ideias, pois seu objetivo é recuperar/reter uma ou algumas palavras-chave.

Embora existam poucos estudos investigando a aplicação desse formato de teste, é possível observar, no entanto, que o TC apresenta uma forma de recuperação simples, possibilitando que a aprendizagem perdue por mais tempo. Esse formato também se mostrou benéfico em diferentes níveis de ensino, desde a Educação Básica até o Ensino Superior (Vojdanoska et al., 2010).

Portanto, sugere-se que testes de complete sejam aplicados com conteúdos que visem a estimular a recordação de uma ou poucas palavras-chaves, termos ou vocabulários, a fim de que permaneçam por mais tempo na memória. A recomendação aos professores é: forneça TC com o objetivo de recuperação de palavras, sem a necessidade de integração do conteúdo todo.

Recordação com pistas (*cued-recall*)

A recordação com pistas (RP) é um formato de teste no qual se fornecem pistas para que se tente lembrar respostas (Lima & Jaeger, 2020). Por exemplo, dada uma questão de complete, fornece-se a primeira letra para facilitar a recordação [e.g. *A palavra sol deriva da palavra latina s_____ (solis)*] (Lima & Jaeger, 2020). Outros exemplos podem ser observados nos estudos de Aslan e Bäuml (2016), que utilizaram testes de RP ao apresentar de duas a quatro letras iniciais de palavras para que os alunos completassem listas de substantivos. Por sua vez, Kliegl et al. (2018) apresentaram como pista versões borradas (pouco nítidas) de fotos para que os alunos as associassem a fotos estudadas inicialmente.

Uma vez que fornece dica para a resposta, a RP facilita a recuperação de informações estudadas anteriormente e aumenta a probabilidade de recuperar informações da memória (Fazio & Agarwal, 2020). Uma das possíveis explicações para essa vantagem consiste em que os benefícios da prática de lembrar são mais robustos quando a recordação inicial é superior a 50% (Rowland, 2014).

Testes de RP já foram utilizados como estratégia de ensino com crianças a partir de dois anos e dez meses (Fritz et al., 2007) de idade. Quanto aos *settings*, pesquisas usando testes de recordação com pistas foram realizadas na escola, em contexto de sala de aula (e.g. Ritchie et al., 2013), e individualmente, em sala separada com o experimentador (mimetizando o *setting* de laboratório) (e.g. Fritz et al., 2007).

Já foram demonstrados os benefícios da RP utilizando-se desde materiais simples — como nomes próprios (Fritz et al., 2007), nomes de conjuntos de categorias taxonômicas (ou seja, categorias que definem grupos de organismos biológicos) e de itens, fotos (com pouca distorção de *pixels*) (Kliegl et al., 2018), mapas fictícios (Ritchie et al., 2013), frases e listas de palavras com seus sinônimos (Goossens et al., 2014a) — até materiais mais complexos — como conjuntos de pares-chave de definição de conceitos (Lipko-Speeda et al., 2014) e textos enciclopédicos (Lima & Jaeger, 2020).

Ritchie et al. (2013) identificaram que testes de RP — realizados a partir de mapas fictícios que apresentavam a localização de algumas cidades como pistas e solicitavam dos estudantes que tentassem lembrar o nome e a sua localização correspondente nos mapas — impulsionaram aprendizagens de longo prazo. Em relação à desvantagem, os testes de RP, quando comparados a testes de múltipla escolha, resultaram em menor recuperação. Ainda assim, ambos os formatos de teste geram melhor desempenho se comparados à condição de releitura do conteúdo (Lima & Jaeger, 2020).

Os benefícios dos testes de RP são mais robustos quando fornecido o *feedback* imediato para crianças em idade pré-escolar (Kliegl et al., 2018). Com relação a alunos universitários, os testes de RP seguidos de *feedback*, comparando-se a condição de prática fácil (em que se apresentam as duas primeiras letras da palavra-alvo) à condição de prática difícil (em que se apresenta apenas a primeira letra da palavra-alvo), mostraram que o desempenho médio de recuperação foi significativamente superior na condição de prática fácil e em curto prazo. No entanto, uma semana depois, o desempenho na condição de prática difícil mostrou-se superior (Kliegl et al., 2018). Assim, os testes de RP permitem variação no nível de dificuldade e, por conseguinte, adaptação das estratégias de prática de lembrar de acordo com a faixa etária e o conhecimento prévio dos estudantes (Fazio & Agarwal, 2020).

Em suma, o ato de apresentar dicas e/ou pistas durante a realização da prática de lembrar pode servir como auxílio, aumentando assim, a probabilidade de que os alunos cheguem às respostas. A recomendação aos professores é: ao fornecer testes de RP, varie o nível de dificuldade de acordo com o conhecimento prévio dos estudantes, cuidando para não facilitar demasiadamente, mas, ao mesmo tempo, impulsionando os alunos a recuperarem informações.

Diretrizes para os educadores

A possibilidade de adaptação e a aplicabilidade da prática de lembrar via diferentes formatos de teste permitem que educadores e estudantes utilizem diferentes estratégias de ensino (orientadas pelo professor) e estudo (realizadas pelos alunos) (Roediger et al., 2011b; Agarwal et al., 2018). Há uma variedade de tipos de questões e formatos de testes que podem ser utilizados em contexto real da sala de aula (McDaniel et al., 2013) de modo a beneficiar as aprendizagens de maneira flexível (McDaniel et al., 2013; Agarwal et al., 2018).

Para facilitar o uso da prática de lembrar, um “guia de prática de lembrar” (Agarwal et al., 2018) e inúmeras pesquisas (e.g. Ekuni & Pompeia, 2020) apresentam sugestões, bem como orientações para que educadores utilizem e incentivem a realização da prática de lembrar no decorrer do processo de ensino e aprendizagem. A partir dos resultados das pesquisas apresentadas nessa revisão, forneceremos um resumo com as principais diretrizes (ver Figura 1).

Ao utilizar a prática de lembrar, é preciso orientar os alunos para que, durante os testes, não consultem materiais, anotações, ou mesmo os colegas (Agarwal et al., 2018). Eles devem ser honestos, buscando na mente informações a que tiveram acesso anteriormente (Agarwal et al., 2018). Uma maneira de incentivar toda a turma a praticar lembrar é ter cautela ao direcionar perguntas. O professor pode, por exemplo, fazer uma pergunta oral, dar certo tempo para que os alunos pensem na resposta e, só então, sortear nomes de alunos para que respondam em voz alta.

Figura 1

Quadro comparativo dos estudos apresentados com os formatos de testes e recomendações para educadores.



Fonte: elaboração dos autores.

Testes de verdadeiro ou falso e de múltipla escolha podem ser abordados em tempo real, a partir de *quizzes* (Agarwal et al., 2013). Uma ideia é usar cartões coloridos (como plaquinhas) para que os alunos levantem a resposta ao comando do professor. Fica ainda melhor se os cartões tiverem cores padronizadas, de modo a manter a correspondência entre as cores e alternativas (Agarwal et al., 2013; Ekuni & Pompeia, 2020) (e.g. o cartão com a alternativa VERDADEIRO pode ser branco, e com a alternativa FALSO, vermelho; com respeito a alternativas de A a D, a orientação é usar cartões de cores diferentes também, como o amarelo para a alternativa A, o azul para B e assim sucessivamente). Dessa maneira, fica fácil de visualizar as alternativas mais escolhidas pelos alunos. Ao fazer uma pergunta (por voz, por projeção em tela, ou por escrito na lousa), deve-se dar tempo para que os alunos pensem na resposta antes de solicitar que levantem os cartões. Para aumentar ainda mais o benefício, deve-se fornecer o *feedback* (e.g. Agarwal et al., 2018; Ekuni & Pompeia, 2020).

Estratégias relacionadas aos testes de resposta curta e recordação livre podem ser postas em prática com bilhetes de entrada e de saída (McDermott et al., 2014; Agarwal et al., 2018). Essa estratégia pode ser realizada usando pedaços de folha de caderno, ou papel sulfite. Conforme os alunos entram na sala de aula, o professor pode entregar os pedaços de papel e pedir, por exemplo, para que os alunos escrevam o que lembram da aula anterior. Com os bilhetes de saída, é possível solicitar que, antes do término da aula, os alunos escrevam, por exemplo, o que acharam mais interessante no assunto abordado (Agarwal et al., 2018).

Ainda considerando os testes de recordação livre, é possível que o professor peça para que os alunos façam um resumo oral ou escrito do conteúdo apresentado anteriormente (Brojde & Wise, 2008). O professor também pode realizar ditados cujo objetivo seja que os alunos recordem a grafia de palavras, significados e definições. Os testes de recordação livre também podem ser estimulados por meio de intervenções durante leituras, nas quais se solicite à criança que aponte imagens no livro (Cornell et al., 1988). Pode-se, além disso, requerer dos alunos que escrevam uma lista com palavras apresentadas anteriormente em um ditado (Jones et al., 2016), ou que escrevam tudo o de que se lembram sobre um texto estudado previamente (Rowley & McCrudden, 2020).

Em relação aos testes de complete, é possível utilizá-los a partir de palavras-alvo destacadas em um texto. No momento do teste, pode-se apresentar a definição para que os estudantes tentem lembrar e completar com a palavra-alvo (Goossens et al., 2014b). Nos testes de complete, pode-se praticar lembrar utilizando pares de palavras associadas, termos e palavras correspondentes etc. (e.g. *arrabalde* e *periferia* – *arrabalde* e _____).

Nos testes de recordação com pistas, podem-se apresentar imagens um tanto distorcidas como dicas e, em seguida, pede-se aos estudantes que as nomeiem (Kliegl et al., 2018). Pode-se estudar frases ou listas de palavras associadas e, no momento da realização do teste, fornecer uma ou duas letras da palavra-alvo para que os estudantes tentem lembrar e escrever a palavra toda (Kliegl et al., 2019).

Portanto, existem, inúmeras possibilidades para diversificar a escolha do formato do teste ao praticar lembrar. É importante considerar que os diferentes tipos de teste beneficiam a aprendizagem (McDermott et al., 2014; Agarwal et al., 2018). Ademais, oferecer o *feedback* após a prática pode ajudar os estudantes a corrigir erros e estimular a retenção de informações corretas (Marsh et al., 2012b). É importante proporcionar testes não apenas com a intenção de avaliar, mas, sim, como estratégia de ensino.

Considerações finais

Tendo em vista uma educação baseada em evidências científicas, a presente revisão aponta diferentes maneiras e formatos de teste para se implementar a prática de lembrar em ambiente de sala de aula. Isso permite contribuir significativamente com aprendizagens duradouras e

flexíveis. Os resultados dos estudos, de modo geral, apresentam que cada formato pode contribuir de maneira efetiva em diferentes conteúdos, materiais e níveis de ensino. Contudo, a presente revisão, por ser narrativa, apresenta limitações quanto à seleção de artigos. Futuros estudos podem realizar uma revisão sistemática sobre o tema, por meio da realização de buscas sistematizadas.

A partir dessa pesquisa, observa-se, ainda, que a idade dos alunos e o nível de ensino parecem ter um papel importante na escolha e na implementação do formato de teste. Com alunos menores, fornecer pistas durante a prática de lembrar pode ser mais eficaz para a aprendizagem. Com estudantes mais velhos, por outro lado, podem-se fornecer testes com uma dificuldade desejável maior. Outro ponto diz respeito ao *feedback*. Quando dado, seja qual for o formato de teste, melhora a retenção e corrige erros. Portanto, os efeitos da prática de lembrar em contexto educacional são robustos. Por essas e outras razões, praticar lembrar é uma promissora estratégia para o ensino e a aprendizagem de estudantes de diferentes faixas etárias.

Referências

- Abbott, E. E. (1909). On the analysis of the factor of recall in the learning process. *Psychological Review: Monograph Supplements*, 11(1), 159-177. <https://doi.org/10.1037/h0093018>
- Agarwal, P. K., & Bain, P. M. (2019). *Powerful Teaching: unleash the science of learning*. Jossey-Bass.
- Agarwal, P. K., D'Antonio, L., Roediger, H. L., McDermott, K. B., & McDaniel, M. A. (2014). Classroom-based programs of retrieval practice reduce middle school and 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., Roediger, H. L., McDaniel, M. A., & McDermott, K. B. (2013). *How to use retrieval practice to improve learning* [E-book]. Washington University in St. Louis. <https://www.retrievalpractice.org/library>
- Agarwal, P. K., Roediger, H. L., McDaniel, M. A., & McDermott, K. B. (2018). *Como a Prática de Lembrar pode ser utilizada para melhorar a aprendizagem* [E-book]. Washington University in St. Louis. <https://www.retrievalpractice.org/library>
- Aslan, A., & Bäuml, K. H. T. (2016). Testing enhances subsequent learning in older but not in younger elementary school children. *Developmental Science*, 19(6), 992-998. <https://doi.org/10.1111/desc.12340>
- Barenberg, J., & Dutke, S. (2019). Testing and metacognition: retrieval practice effects on metacognitive monitoring in learning from text. *Memory*, 27(3), 269-279. <https://doi.org/10.1080/09658211.2018.1506481>
- Barenberg, J., Berse, T., Reimann, L., & Dutke, S. (2021). Testing and transfer: Retrieval practice effects across test formats in English vocabulary learning in school. *Applied Cognitive Psychology*, 35(3), 700-710. <https://doi.org/10.1002/acp.3796>
- Bjork, R. A. (1975). Retrieval as a Memory Modifier: an interpretation of negative recency and related phenomena. Em R. L. Solso (Ed.). *Information processing and cognition: The Loyola Symposium* (pp. 123-144). Lawrence Erlbaum Associates.
- Brabec, J. A., Pan, S. C., Bjork, E. L., & Bjork, R. A. (2021). True-False Testing on Trial: Guilty as

- Charged or Falsely Accused? *Educational Psychology Review*, 33(2), 667-692.
<https://doi.org/10.1007/s10648-020-09546-w>
- Brojde, C. L., & Wise, B. W. (2008). An Evaluation of the Testing Effect with Third Grade Students. *Proceedings of the 3-Th Annual Meeting of the Cognitive Science Society*, 1362-1367.
- Butler, A. C. (2018). Multiple-Choice Testing in Education: Are the Best Practices for Assessment Also Good for Learning? *Journal of Applied Research in Memory and Cognition*, 7(3), 323-331. <https://doi.org/10.1016/j.jarmac.2018.07.002>
- Butler, A. C., & Roediger, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, 19(4-5), 514-527.
<https://doi.org/10.1080/09541440701326097>
- Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory and Cognition*, 36(3), 604-616.
<https://doi.org/10.3758/MC.36.3.604>
- Butler, A. C., Marsh, E. J., Goode, M. K., & Roediger, H. L. (2006). When additional multiple-choice lures aid versus hinder later memory. *Applied Cognitive Psychology*, 20(7), 941-956.
<https://doi.org/10.1002/acp.1239>
- Cantor, A. D., Eslick, A. N., Marsh, E. J., Bjork, R. A., & Bjork, E. L. (2015). Multiple-choice tests stabilize access to marginal knowledge. *Memory & Cognition*, 43(2), 193-205.
<https://doi.org/10.3758/s13421-014-0462-6>
- Carpenter, S. K. (2011). Semantic Information Activated During Retrieval Contributes to Later Retention: Support for the Mediator Effectiveness Hypothesis of the Testing Effect. *Journal of Experimental Psychology: Learning Memory and Cognition*, 37(6), 1547-1552.
<https://doi.org/10.1037/a0024140>
- Carpenter, S. K., Lund, T. J. S., Coffman, C. R., Armstrong, P. I., Lamm, M. H., & Reason, R. D. (2016). A Classroom Study on the Relationship Between Student Achievement and Retrieval-Enhanced Learning. *Educational Psychology Review*, 28(2), 353-375.
<https://doi.org/10.1007/s10648-015-9311-9>
- Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U.S. history facts. *Applied Cognitive Psychology*, 23(6), 760-771.
<https://doi.org/10.1002/acp.1507>
- Collins, A. J., & Fauser, C. J. M. B. (2005). Balancing the strengths of systematic and narrative reviews. *Human Reproduction Update*, 11 (2), 103-104.
<https://doi.org/10.1093/humupd/dmh058>
- Cornell, E. H., Sénéchal, M., & Broda, L. S. (1988). Recall of Picture Books by 3-Year-Old Children: Testing and Repetition Effects in Joint Reading Activities. *Journal of Educational Psychology*, 80(4), 537-542. <https://doi.org/10.1037/0022-0663.80.4.537>
- 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>
- Dirkx, K. J. H., Kester, L., & Kirschner, P. A. (2014). The testing effect for learning principles and procedures from texts. *Journal of Educational Research*, 107(5), 357-364.
<https://doi.org/10.1080/00220671.2013.823370>
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, Supplement*, 14(1), 4-58. <https://doi.org/10.1177/1529100612453266>
- Ekuni, R., & Pompeia, S. (2020). Prática De Lembrar: a Quais Fatores Os Educadores Devem Se Atentar? *Psicologia Escolar e Educacional*, 24, e220284. <https://doi.org/10.1590/2175->

35392020220284

- Ekuni, R., Souza, B. M. N., Agarwal, P. K., & Pompeia, S. (2020). A conceptual replication of survey research on study strategies in a diverse, non- WEIRD student population. *Scholarship of Teaching and Learning in Psychology*, Advance online publication, 8(1), 1-14. <https://doi.org/10.1037/stl0000191>
- Elias, C. D. S. R., da Silva, L. A., Martins, M. T. D. S. L., Ramos, N. A. P., de Souza, M. D. G. G., & Hipólito, R. L. (2012). Quando chega o fim? Uma revisão narrativa sobre terminalidade do período escolar para alunos deficientes mentais. *SMAD, Revista Electrónica en Salud Mental, Alcohol y Drogas*, 8(1), 48-53. <https://doi.org/10.11606/issn.1806-6976.v8i1p48-53>
- Enders, N., Gaschler, R., & Kubik, V. (2021). Online Quizzes with Closed Questions in Formal Assessment: How Elaborate Feedback can Promote Learning. *Psychology Learning and Teaching*, 20(1), 91-106. <https://doi.org/10.1177/1475725720971205>
- Endres, T., Kranzdorf, L., Schneider, V., & Renkl, A. (2020). It matters how to recall – task differences in retrieval practice. *Instructional Science*, 48(6), 699-728. <https://doi.org/10.1007/s11251-020-09526-1>
- Fazio, L. K., & Agarwal, P. K. (2020). *How To Implement Retrieval-Based Learning in Early Childhood* [E-book]. Washington University in St. Louis. <https://www.retrievalpractice.org/library>
- Foss, D. J., & Pirozzolo, J. W. (2017). Four semesters investigating frequency of testing, the testing effect, and transfer of training. *Journal of Educational Psychology*, 109(8), 1067-1083. <https://doi.org/10.1037/edu0000197>
- 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. (2014a). The effect of retrieval practice in primary school vocabulary learning. *Applied Cognitive Psychology*, 28(1), 135-142. <https://doi.org/10.1002/acp.2956>
- Goossens, N. A. M. C., Camp, G., Verkoeijen, P. P. J. L., Tabbers, H. K., & Zwaan, R. A. (2014b). The benefit of retrieval practice over elaborative restudy in primary school vocabulary learning. *Journal of Applied Research in Memory and Cognition*, 3(3), 177-182. <https://doi.org/10.1016/j.jarmac.2014.05.003>
- 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>
- Greving, S., & Richter, T. (2018). Examining the testing effect in university teaching: Retrievability and question format matter. *Frontiers in Psychology*, 9, 02412. <https://doi.org/10.3389/fpsyg.2018.02412>
- Hinze, S. R., & Wiley, J. (2011). Testing the limits of testing effects using completion tests. *Memory*, 19(3), 290-304. <https://doi.org/10.1080/09658211.2011.560121>
- Jaeger, A., Eisenkraemer, R. E., & Stein, L. M. (2015). Test-enhanced learning in third-grade children. *Educational Psychology*, 35(4), 513-521. <https://doi.org/10.1080/01443410.2014.963030>
- Jones, A. C., Wardlow, L., Pan, S. C., Zepeda, C., Heyman, G. D., Dunlosky, J., & Rickard, T. C. (2016). Beyond the Rainbow: Retrieval Practice Leads to Better Spelling than does Rainbow Writing. *Educational Psychology Review*, 28(2), 385-400. <https://doi.org/10.1007/s10648-015-9330-6>
- 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., & 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., Blunt, J. R., Smith, M. A., & Karpicke, S. S. (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/10.1016/j.jarmac.2014.07.008>
- 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>
- Kliegl, O., Abel, M., & Bäuml, K. H. T. (2018). A (preliminary) recipe for obtaining a testing effect in preschool children: Two critical ingredients. *Frontiers in Psychology*, 9, 01446. <https://doi.org/10.3389/fpsyg.2018.01446>
- Kliegl, O., Bjork, R. A., & Bäuml, K. H. T. (2019). Feedback at test can reverse the retrieval-effort effect. *Frontiers in Psychology*, 10, 1863. <https://doi.org/10.3389/fpsyg.2019.01863>
- Larsen, D. P., Butler, A. C., & Roediger, H. L. (2008). Test-enhanced learning in medical education. *Medical Education*, 42(10), 959-966. <https://doi.org/10.1111/j.1365-2923.2008.03124.x>
- Larsen, D. P., Butler, A. C., & Roediger, H. L. (2009). Repeated testing improves long-term retention relative to repeated study: A randomised controlled trial. *Medical Education*, 43(12), 1174-1181. <https://doi.org/10.1111/j.1365-2923.2009.03518.x>
- Lima, N. K. de, & Jaeger, A. (2020). The Effects of Prequestions versus Postquestions on Memory Retention in Children. *Journal of Applied Research in Memory and Cognition*, 9(4), 555-563. <https://doi.org/10.1016/j.jarmac.2020.08.005>
- Lipko-Speeda, 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>
- Lipowski, S. L., Pyc, M. A., Dunlosky, J., & Rawson, K. A. (2014). Establishing and explaining the testing effect in free recall for young children. *Developmental Psychology*, 50(4), 994-1000. <https://doi.org/10.1037/a0035202>
- 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>
- Lyle, K. B., & Crawford, N. A. (2011). Retrieving Essential Material at the End of Lectures Improves Performance on Statistics Exams. *Teaching of Psychology*, 38(2), 94-97. <https://doi.org/10.1177/0098628311401587>
- Marsh, E. J., Fazio, L. K., & Goswick, A. E. (2012b). Memorial consequences of testing school-aged children. *Memory*, 20(8), 899-906. <https://doi.org/10.1080/09658211.2012.708757>
- Marsh, E. J., Lozito, J. P., Umanath, S., Bjork, E. L., & Bjork, R. A. (2012a). Using verification feedback to correct errors made on a multiple-choice test. *Memory*, 20(6), 645-653. <https://doi.org/10.1080/09658211.2012.684882>
- Marsh, E. J., Roediger, H. L., Bjork, R. A., & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin & Review*, 14(2), 194-199. <https://doi.org/10.3758/BF03194051>
- McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., & Roediger, H. L. (2013). Quizzing in Middle-School Science: Successful Transfer Performance on Classroom Exams. *Applied Cognitive Psychology*, 27(3), 360-372. <https://doi.org/10.1002/acp.2914>
- McDermott, K. B. (2021). Practicing Retrieval Facilitates Learning. *Annual Review of Psychology*, 72, 609-633. <https://doi.org/10.1146/annurev-psych-010419-051019>

- McDermott, K. B., Agarwal, P. K., D'Antonio, L., Roediger, H. L. I., & 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>
- Moreira, B. F. T., Pinto, T. S. da S., Justi, F. R. R., & Jaeger, A. (2019). Retrieval practice improves learning in children with diverse visual word recognition skills. *Memory*, 27(10), 1423-1437. <https://doi.org/10.1080/09658211.2019.1668017>
- Ritchie, S. J., Della Sala, S., & McIntosh, R. D. (2013). Retrieval practice, with or without mind mapping, boosts fact learning in primary school children. *PLoS ONE*, 8(11), e78976. <https://doi.org/10.1371/journal.pone.0078976>
- 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. Em A. S. Benjamin (Ed.). *Successful Remembering and Successful Forgetting: a festschrift in honor of Robert A. Bjork*. (pp. 70-115). Taylor and Francis Group.
- 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., & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition*, 1(4), 242-248. <https://doi.org/10.1016/j.jarmac.2012.09.002>
- Roediger, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011b). 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., Putnam, A. L., & Smith, M. A. (2011a). Ten benefits of testing and their applications to educational practice. Em J. P. Mestre & B. R. Ross. *The psychology of learning and motivation: Cognition in education* (pp. 1-36). Elsevier Academic Press.
- 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>
- Rother, E. T. (2007). Revisión sistemática X Revisión narrativa. *Acta paulista de enfermagem*, 20, v-vi. <https://doi.org/10.1590/S0103-21002007000200001>
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432-1463. <https://doi.org/10.1037/a0037559>
- Rowley, T., & McCrudden, M. T. (2020). Retrieval practice and retention of course content in a middle school science classroom. *Applied Cognitive Psychology*, 34(6), 1510-1515. <https://doi.org/10.1002/acp.3710>
- Santrock, J. W. (2009). *Psicologia Educacional*. McGraw-Hill.
- Schaap, L., Verkoeijen, P., & Schmidt, H. (2014). Effects of different types of true-false questions on memory awareness and long-term retention. *Assessment and Evaluation in Higher Education*, 39(5), 625-640. <https://doi.org/10.1080/02602938.2013.860422>
- 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>

- Stenlund, T., Sundström, A., & Jonsson, B. (2016). Effects of repeated testing on short- and long-term memory performance across different test formats. *Educational Psychology, 36*(10), 1710-1727. <https://doi.org/10.1080/01443410.2014.953037>
- Tulving, E. (1967). The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior, 6*(2), 175-184. [https://doi.org/10.1016/S0022-5371\(67\)80092-6](https://doi.org/10.1016/S0022-5371(67)80092-6)
- Uner, O., Tekin, E., Roediger, H. L. (2022). True–false tests enhance retention relative to rereading. *Journal of Experimental Psychology: Applied, 28*(1), 114-129. <https://doi.org/10.1037/xap0000363>
- Van den Broek, G., Takashima, A., Wiklund-Hörnqvist, C., Wirebring, L. K., 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>
- Vojdanoska, M., Cranney, J., & Newell, B. R. (2010). The testing effect: The role of feedback and collaboration in a tertiary classroom setting. *Applied Cognitive Psychology, 24*(8), 1183-1195. <https://doi.org/10.1002/acp.1630>
- Wiklund-Hörnqvist, C., Jonsson, B., & Nyberg, L. (2014). Strengthening concept learning by repeated testing. *Scandinavian Journal of Psychology, 55*(1), 10-16. <https://doi.org/10.1111/sjop.12093>
- Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin, 147*(4), 399-435. <https://doi.org/10.1037/bul0000309>