

# **A importância da prática distribuída na aprendizagem: efeito de espaçamento e efeito de teste**

**Álvaro Júnio Machado da Costa Rios**

Dissertação para obtenção do Grau de Mestre em  
**Medicina**  
(mestrado integrado)

Orientador: Prof. Doutor Miguel Castelo–Branco Sousa

**maio de 2021**



# **Agradecimentos**

Ao meu orientador, Prof. Doutor Miguel Castelo Branco pela ajuda e disponibilidade.

Aos meus pais, ao meu irmão e à Camila pelo apoio incondicional.

Aos amigos que levo para a vida.

À Covilhã, pelas memórias.



# Resumo

**Introdução:** As primeiras investigações sobre os temas da memória e esquecimento, surgiram no ano de 1880, elaboradas pelo psicólogo *Hermann Ebbinghaus*, com a denominada curva de esquecimento. Foi a partir deste trabalho que se desenvolveram métodos de investigação com o intuito de atenuar esta curva.

Durante o processo de estudo, evidenciaram-se dois efeitos: o efeito do espaçamento e o efeito de teste. O primeiro, refere-se à melhoria na aprendizagem e da memória, quando os episódios de estudo são distribuídos no tempo, ao invés de quando sucedem num único episódio. Já o segundo, é baseado na evidência de que os alunos quando são testados numa certa matéria, conseguem recordar melhor, do que quando não são testados. Aliados a estas evidências existem outras variáveis a ter em consideração, uma vez que podem ter influência na qualidade destes efeitos.

**Objetivo:** A presente revisão bibliográfica pretende reunir e rever as informações existentes sobre o efeito de espaçamento e do efeito do teste, de forma a identificar as variáveis que mais os influenciam e, assim, demarcar as condições necessárias para otimizar os efeitos referidos.

**Metodologia:** Para a realização desta dissertação foi realizada pesquisa bibliográfica e análise de artigos indexados na base de dados Pubmed e Google Scholar até ao ano de 2020, utilizando-se palavras-chave “forgetting curve”, “spaced repetition”, “spaced education”, “spacing effect” e “testing effect”.

**Conclusões:** Ambos os efeitos, apresentam evidências robustas na literatura, sobre a sua efetividade, em diferentes matérias de aprendizagem, como em diferentes faixas etárias e contextos.

O efeito de espaçamento é máximo quando o intervalo interestudo é cerca de 10%-20% do intervalo de retenção – intervalo de tempo que vai desde a última sessão de estudo até ao teste de avaliação.

Relativamente ao efeito de teste, tende a ser máximo quando, na fase inicial de aprendizagem, se utilizam perguntas de resposta curta, invés de perguntas de escolha múltipla, para a recuperação da informação estudada. Existem benefícios, quando as perguntas dos testes fornecem *feedback* explicativo – informação adicional a explicar o raciocínio. É preferível fazer recuperações da matéria, mais difíceis, mas bem-sucedidas, do que mais fáceis. Os alunos adotam os testes mais facilmente como

método de aprendizagem, quando os docentes lhes mostram, explicitamente, o efeito de teste. Por fim, a abordagem mais custo-efetiva, passa por fazer uma recuperação bem sucedida, na fase inicial da aprendizagem, seguida de 3 ou 4 testes subsequentes, de preferência aumentando o período de tempo entre cada um deles.

## **Palavras-chave**

forgetting curve;spaced repetition;spaced education;spacing effect;testing effect

# Abstract

**Introduction:** The first researches regarding topics of memory and forgetfulness were made in the year 1880, by the psychologist Hermann Ebbinghaus, with the called forgetting curve. It was based on this work that research methods were developed in order to reduce this curve.

During the study process, two effects were highlighted: the spacing effect and the testing effect. The first one refers to the improvement in the learning process and memory, when the study moments are distributed over time, instead of when they occur in a single moment. The second one, on the other hand, is based on the evidence that when students are tested in a certain subject, they can retain more of what they studied than when they are not tested. In addition to this evidence, there are other variables to be taken into account, since they can influence the quality of these effects.

**Objective:** The present bibliographic review intends to gather and review the existing information about the spacing effect and the testing effect, in order to identify the variables that most influence them and, consequently, distinguish the necessary conditions to optimize the referred effects.

**Methods:** In order to develop this thesis, a bibliographic search was made, as well as the analysis of articles indexed in the PubMed and Google Scholar database until the year 2020, using the keywords “forgetting curve”, “spaced repetition”, “spaced education”, “Spacing effect”, “testing effect” etc.

**Conclusions:** Both effects present strong evidence in the literature about its effectiveness, in different learning subjects, as well as in different age groups and contexts.

The spacing effect is maximum when the study interval is about 10% -20% of the retention interval – This time lapse is from the last study session to the evaluation test. Regarding the testing effect, it tends to be maximum when, during the initial learning phase, questions with a small answer are used, instead of multiple-choice questions, for the retrieval of the studied information. There are benefits when test questions provide explanatory feedback - additional information to explain the reasoning. It's preferable to make recoveries of matter, more difficult, but successful, than easier. Students adopt tests more easily as a learning method, when teachers explicitly show them the test effect. Finally, the most cost-effective approach is to make a successful recovery, in the

initial phase of learning, followed by 3 or 4 subsequent tests, preferably increasing the time between each one.

## **Keywords**

forgetting curve; spaced repetition; spaced education; spacing effect; testing effect.

# Índice

1	Introdução	1
1.1	Metodologia	1
2	Memória	3
2.1	O que é a memória?	3
2.2	Memória de curto-prazo vs. longo-prazo	3
3	Curva de esquecimento	5
4	Prática Distribuída	7
4.1	Efeito do espaçamento	7
4.1.1	Estudo espaçado versus Estudo massivo	8
4.1.2	<i>Timing</i>	9
4.1.3	Espaçamento igual ou sucessivamente maior	11
4.1.4	Tipo de materiais	11
4.1.5	Características dos alunos/Abrangência	12
4.2	Efeitos de teste	12
4.2.1	Efeitos de teste versus reestudar	13
4.2.2	Efeitos primários do efeito de teste	14
4.2.2.1	Características dos alunos/Abrangência	14
4.2.2.2	Tipos de materiais	14
4.2.3	Efeitos secundários ao efeito de teste	14
4.2.4	Formas de potenciar o efeito de teste	15
4.2.4.1	Formato de perguntas de teste	15
4.2.4.2	<i>Feedback</i> das perguntas	16
4.2.4.3	Tornar o efeito de teste explícito	19
4.2.4.4	Cronograma de teste	20
4.2.4.5	Hipótese do esforço da recuperação	20
5	Exemplos da aplicabilidade na medicina	23
6	Problemas de implementação	29
7	Possíveis soluções para implementação	31
8	Conclusões	33
	Referências bibliográficas	35



# Lista de Figuras

Figura 1 – Curva de Esquecimento de <i>Ebbinghaus</i> ;	5
Figura 2 – Taxa de retenção ao longo do tempo com a repetições espaçadas – exemplo hipotético do impacto ao longo do tempo na curva de retenção;	6
Figura 3 – Procedimento experimental típico para o estudo do efeito de espaçamento, comparando com o estudo massivo;	7
Figura 4 - Design básico do estudo do efeito de espaçamento;	8
Figura 5 - Relação entre o rácio IIE/IR e a percentagem correta de recuperação de informações;	10
Figura 6 - Proporção de respostas corretas no teste de resposta curta final, em função da condição de aprendizagem pós-aula (as barras de erro apresentam um intervalo de confiança de 95%);	16
Figura 7 – Proporção de respostas corretas no teste final, em função da confiança de resposta e da condição de aprendizagem para respostas que estavam incorretas (lado esquerdo) e corretas (lado direito) no teste de escolha múltipla inicial (MC);	19
Figura 8 - Estrutura do ensaio clínico randomizado;	23
Figura 9 - Curvas de esquecimento para (A) ensino através da Web; (B) estudo espaçado; (C) ambos;	24
Figura 10 - Esquerda: todos os alunos aumentaram significativamente a velocidade de sutura e do nó – independentemente do método de treino. Direita: 3 horas de treino laparoscópico não afeta a precisão, significativamente;	25
Figura 11 - Antes do treino, ambos os grupos tinham desempenhos quase semelhantes (esquerda) e pontuações de qualidade de nó (direita). A aprendizagem espaçada foi significativamente mais efetiva; apenas após três sessões espaçadas de 40 minutos (total de 120 minutos) comparativamente com 180 minutos de treino sem interrupções;	25
Figura 12 – No grupo do treino espaçado, o grau de força de sutura foi mais do que duplicada após 3 horas de treino, enquanto o grupo controlo não mostraram nenhuma melhoria significativa;	26
Figura 13 - Impacto da educação espaçada na triagem de PSA inadequada. As percentagens de rastreio de PSA inadequada durante e após a intervenção da educação espaçada são representadas pelos triângulos cinza e preto, respetivamente;	27
Figura 14 - Fluxograma do curso.	28



## **Lista de Acrónimos**

IIE	Intervalo interestudo
IR	Intervalo de retenção
MCP	Memória curto-prazo
MLP	Memória longo-prazo



# **1. Introdução**

## **1.1. Metodologia**

Realizou-se uma revisão da literatura científica sobre o tema, através da base de dados eletrônica PubMed e Google Scholar, no período de novembro de 2020 a fevereiro de 2021. Para efetuar a pesquisa utilizaram-se as palavras-chave “forggetting curve”, “spaced repetition”, “spaced education”, “spacing Effect”, “testing effect”. Foram incluídos todos os artigos relevantes, incluindo artigos científicos, artigos de revisão e estudos randomizados, publicados até dia 31 de dezembro de 2020, privilegiando os mais recentes. Quando encontradas referências bibliográficas pertinentes nos artigos previamente selecionados foi feita a pesquisa e consulta do artigo original. Filtraram-se os artigos pelo idioma, incluindo na pesquisa aqueles publicados em inglês e português.



## **2. Memória**

### **2.1. O que é a memória?**

De forma genérica, memória é a capacidade de adquirir, armazenar e recuperar informações disponíveis no cérebro. A memória não é uma única faculdade da mente, mas sim formada por sistemas que têm vários princípios operacionais e diferentes neuroanatomias [1]. A memória é formada pela criação e ligação de novos neurónios. Uma memória específica é criada pelo padrão, forma e tipos de ligações dos neurónios, nos diferentes aglomerados dentro do cérebro. Ao longo do tempo, esses padrões podem ser transferidos de zonas de memória de curto-prazo - associadas mais fortemente ao córtex pré-frontal - para regiões de memória de longo prazo. Essa transferência pode ocorrer através do processo de re-exposição e de recuperação. Ao recuperar memórias específicas, estamos a adicionar sinapses neuronais ao aglomerado de neurónios dessa memória e, com isto, podemos mudar, atualizar e reforçar a mesma. Este processo é conhecido como neuroplasticidade. Inicialmente os neurónios transcrevem memórias no córtex a curto-prazo. A recuperação ajuda a adicionar mais sinapses neuronais e fortalecer a memória para o armazenamento de longo-prazo [2].

### **2.2. Memória de curto-prazo vs. longo-prazo**

A memória de curto-prazo (MCP), também conhecida como memória primária ou ativa, é responsável pela retenção de pequenas porções de informação, durante um curto período de tempo, geralmente 30 segundos. Em contrapartida, a memória de longo-prazo (MLP) pode possuir uma quantidade indefinida de informação. Os dois sistemas estão estreitamente ligados, embora a grande diferença entre as duas seja a nível funcional. Na prática, a MCP funciona como um “caderno de rascunho”, para recordar um número limitado de informação, temporariamente, prontas para serem processadas através da atenção e da recuperação. Em contrapartida, as informações no armazenamento da MLP consistem em memórias para a realização de ações ou capacidades - isto é, memória procedimentais, “*knowing how*” - e memórias de fatos, conceitos e eventos - ou seja, memórias declarativas, “*knowing that*” [3].



### 3. Curva de esquecimento

As primeiras investigações efetuadas na área da memória e da aprendizagem espaçada, foram realizadas por *Hermann Ebbinghaus*. A “curva de esquecimento” ou “*The 1880 Ebbinghaus curve*” (Figura 1) demonstra que a retenção da memória diminui ao longo do tempo. Sugere que informações e memórias relevantes são perdidas com o passar do tempo, quando não existe um esforço para as reter. A “curva de esquecimento” elaborada por Ebbinghaus aponta que os intervenientes tendem a esquecer mais de 50% da informação recentemente aprendida, 20 minutos após o fim da aprendizagem. Para além disso, caso não exista a recuperação ou reexposição ao que foi aprendido, ocorre um decréscimo da taxa de retenção para os 40%, passadas 9 horas, ou para os 24% ao fim de 31 dias [4].

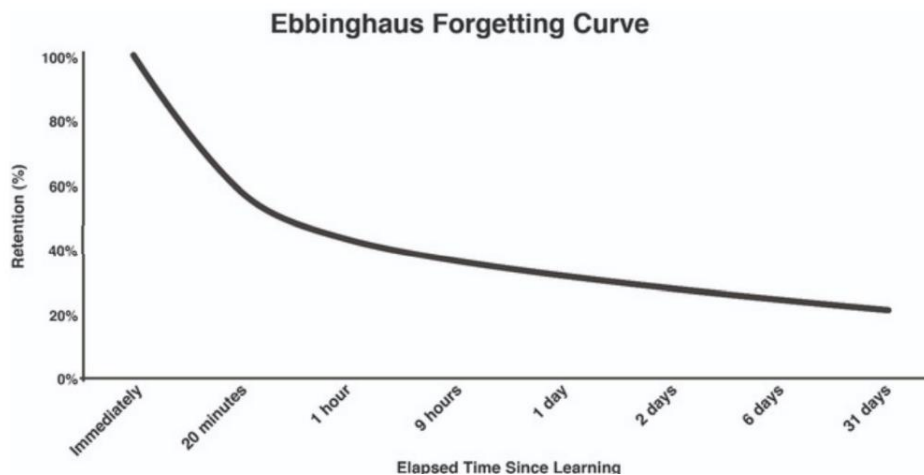


Figura 1 – Curva de Esquecimento de Ebbinghaus [4].

A exposição repetida ao material de aprendizagem, ajudará os intervenientes a reter o conhecimento durante períodos mais prolongados. As reativações da memória fortalecem a memória de longo-prazo ao iniciar a reconsolidação sináptica [5] [6].

Centenas de estudos na psicologia cognitiva e educacional demonstraram que, o espaçamento de encontros repetidos com o material de aprendizagem ao longo do tempo, produz uma aprendizagem superior a nível da memória de longo-prazo. Com estas intervenções espaçadas é possível atenuar, consecutivamente, a curva de esquecimento, atingindo um platô, ao fim de algumas reexposições (Figura 2) [7].

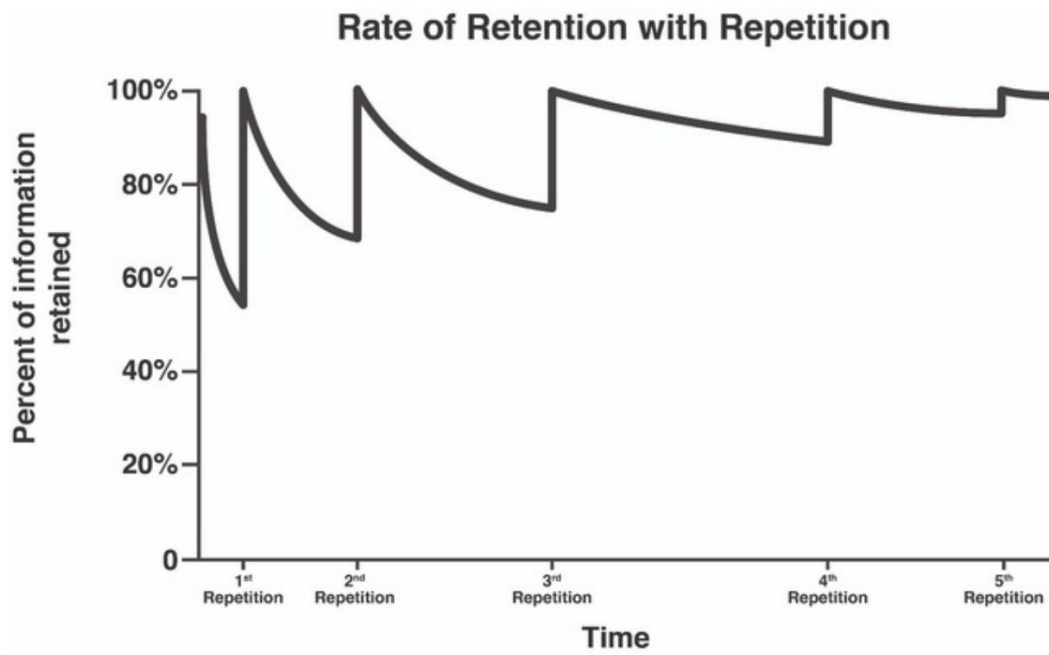


Figura 2 – Taxa de retenção ao longo do tempo com a repetições espaçadas – exemplo hipotético do impacto ao longo do tempo na curva de retenção [7].

## 4. Prática distribuída

O termo Prática distribuída ou do inglês, “distributed practice”, refere-se a um cronograma própria, de dois ou mais episódios de aprendizagem, em contraste com um tipo específico de aprendizagem que se resume a um único episódio. Desta concepção, nasce o “efeito da prática distribuída”, através do qual foi possível evidenciar uma melhoria na aprendizagem quando os episódios são espaçados no tempo, do que quando acontecem apenas num único episódio [8] [9]. Esta melhoria na aprendizagem pode ocorrer com recurso a dois efeitos: efeito de espaçamento ou pelo efeito de teste, que serão discutidos posteriormente.

### 4.1. Efeito do espaçamento

Vários estudos demonstram que estudar uma nova informação ao longo de duas ou mais sessões, espaçadas no tempo, leva a uma maior retenção a longo-prazo, do que passar a mesma quantidade de tempo a estudar a mesma informação numa única sessão (Figura 3).

Quando a interrupção espaçada entre duas ou mais apresentações da mesma informação é zero (por exemplo, é-nos apresentado o mesmo conceito e definição biológica consecutivamente, sem nenhuma interrupção entre eles), dizemos que a apresentação da nova informação é massiva ou agrupada. Já quando o intervalo entre as apresentações é maior do que zero (por exemplo, quando um conceito biológico é repetido a cada 5 minutos ou após a apresentação de 5 diferentes conceitos), do inglês “*spacing gap*”, estamos perante uma apresentação espaçada ou distribuída, uma vez que a informação é separada por um intervalo de tempo diferente de zero [10].

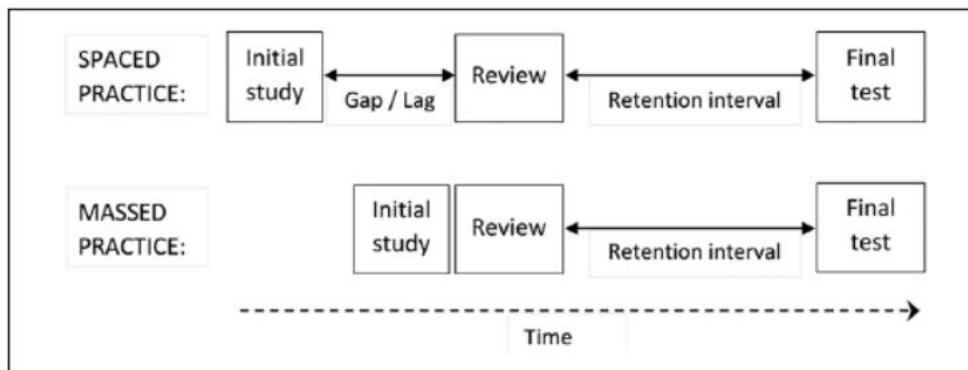


Figura 3 – Procedimento experimental típico para o estudo do efeito de espaçamento, comparando com o estudo massivo [10].

Este intervalo pode variar entre alguns segundos a várias semanas [11].

A este fenómeno, chamamos de “efeito de espaçamento” ou, do inglês, “spacing effect”. As primeiras descobertas foram realizadas por Ebbinghaus, que demonstrou há mais de 100 anos as

primeiras evidências deste fenómeno, depois disso, foram publicados centenas de estudos a explicar os benefícios do espaçamento [12].

#### 4.1.1. Estudo espaçado versus Estudo massivo

Ao longo dos últimos anos foi possível concluir que o efeito de espaçamento é robusta. Cepeda *et al*, em 2006, reviram 254 estudos que envolveram mais de 14000 participantes no total. Concluíram que, os alunos recordavam mais informação após estudo espaçado, do que após estudo massivo, cerca de 47% contra 37%, para informações verbais – por exemplo, palavras, frases, fatos e citações. Apenas 12 dos 271 estudos que comparavam o desempenho entre estudo espaçado e o estudo massivo, mostravam não haver efeito ou até mesmo um efeito negativo no desempenho de quem tinha um estudo espaçado [12].

Utilizando o design básico utilizado nestes tipos de estudo (Figura 4), o aluno estuda a mesma informação em duas sessões, separadas por um intervalo de tempo, designado de intervalo interestudo – IIE. Posteriormente, fará um teste final de avaliação. O intervalo de tempo que separa a última sessão de estudo do teste final designa-se de intervalo de retenção – IR [13].

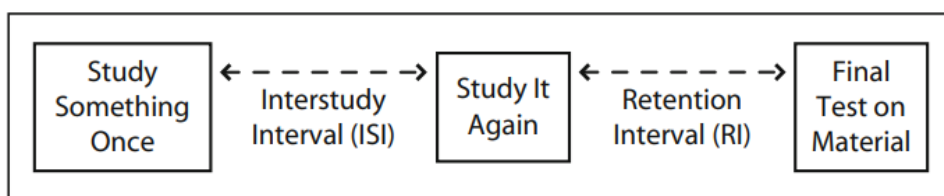


Figura 4 – Design básico do estudo do efeito de espaçamento [13].

Foi analisado a interação entre a magnitude do efeito de espaçamento e o IR, calculando a diferença de desempenho entre as sessões em massas e espaçadas, para cada um dos sete IRs - que variaram entre 1 a 59 segundos até 30 ou mais dias. No âmbito geral, não existem indícios em que o estudo massivo seja preferível ao espaçado, tanto para IR muito curtos - menos de 1 minuto - ou para intervalos mais longos -mais de 30 dias -. Assim, parece existir maior benefício na retenção de informação quando esta é estudada em duas sessões do que em apenas uma. De realçar, que em todos os estudos analisados nesta revisão, a quantidade de tempo de estudo para cada item de informação foi igual, tanto para a forma espaçada como para a massiva. Assim, podemos descartar que, o efeito de espaçamento não se deve ao tempo gasto em cada item de informação [12].

Por outro lado, enquanto o efeito do espaçamento tem um impacto maior na recuperação da informação a longo prazo, o estudo massivo parece mais efetivo do que o espaçado a curto-prazo. Num estudo de 2005, em duas experiências, 423 estudantes universitários leram um texto uma única vez (*single*), duas vezes de forma agrupada/massiva (*massive*), ou duas vezes com uma semana de intervalo entre cada (*distributed*). Os alunos foram testados imediatamente e 2 dias após o estudo. No teste imediato, o desempenho foi maior após a leitura massiva do que a leitura única, enquanto a leitura distribuída não foi, significativamente, melhor do que a leitura única. No

teste 2 dias após o estudo, o desempenho foi melhor após a leitura distribuída do que a leitura única, enquanto houve uma diferença significativa entre a releitura massiva em relação à leitura única. Assim, a releitura em massa melhora o desempenho a curto-prazo, num teste imediato. No entanto, a releitura espaçada tem melhor desempenho a longo-prazo, num teste 2 dias após o estudo [14].

#### **4.1.2. Timing**

Uma das variáveis com mais impacto no efeito do estudo espaçado é o *timing*, isto é, saber qual a melhor oportunidade para estudar, de forma a potenciar o efeito ao máximo.

Num estudo de 2008, após a análise da literatura, concluiu-se que, a maioria dos estudos existentes envolviam IIE curtos. Epilogavam que o efeito de variar o IIE tinha uma forma não linear, isto é, o desempenho num teste final aumentava à medida que o IIE aumentava a partir de zero e, depois diminuía a partir de um valor ótimo. De forma a poder generalizar estes resultados para outros contextos, realizou-se estudos com matérias que poderiam ser representativas das pequenas tarefas do dia a dia, que a maioria das pessoas estão sujeitas na sua vida, tais como factos ou vocabulário [15] [16].

Num desses estudos, utilizou-se como referência um IR de 10 dias e pares de palavras *Swahili*-inglês. Na sessão 1, os alunos aprenderam pares de palavras traduzidas -em cada tentativa, aparecia a palavra *swahili* e o aluno digitava a palavra correspondente em inglês, recebendo *feedback* de imediato. Na sessão 2, um número fixo de pares foram, novamente, mostrados aos alunos. O IIE de 15min a 1 dia, aumentou o desempenho no teste final. Contudo, quando o IIE aumenta para além do 1 dia, o desempenho do teste final, diminui em pequena quantidade [16].

Num outro estudo, utilizou-se intervalos de retenção (IR) de 6 meses. Aqui, o desempenho no teste final aumentou até que o IIE atingiu cerca de 1 mês, a partir do qual voltou a diminuir. Assim, o IIE ótimo parece aumentar à medida que o IR aumenta. Todavia, parece que quando se trata de intervalos de retenção IR significativamente longos, a razão IIE:IR não é de 1:1, tal como sugeriu *Crowder*, em 1976, mas antes o IIE ser cerca de 10%-20% do RI [13].

Para verificar estas conclusões, uma nova experiência, com recolha via web, utilizou um IIE que variava de 20 minutos a 15 semanas e um IR que variava de uma a 50 semanas. Os resultados até ao momento (cerca de 1800 pessoas) sugere que quando o IR é de 1 semana, o IIE ótimo é de 1 dia. Mas, para um IR de 50 semanas, o IIE de 3 semanas é o valor com o melhor desempenho no teste final [13].

Assim, o IIE ideal aumenta à medida que o IR aumenta.

Com uma meta-análise dos efeitos da prática distribuída, baseado em 317 experiências de recuperação verbal de 184 artigos diferentes, foi possível confirmar a relação entre o rácio IIE/IR no desempenho de um teste final, como tratando-se de uma função “*nonmonotonic*”, em forma de

U invertido [12] [17] [18] [19]. Assim, a curva tende a ser um pouco assimétrica, com uma inclinação mais acentuada à medida que o desempenho aumenta para o pico máximo e, depois, uma inclinação mais rasa à medida que o desempenho diminui com IIE mais longos. Isto, pode ser demonstrado na Figura 5, em resultado da experiência de *Cepeda et al 2008* [15], em que o IIE e o IR foram variados durante vários dias [13].

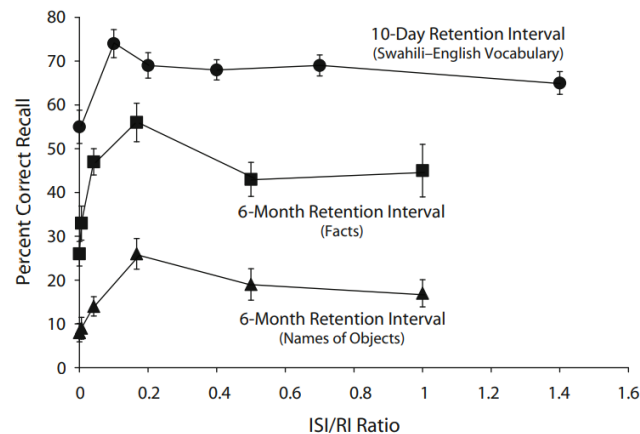


Figura 5 – Relação entre o rácio IIE/IR e a percentagem correta de recuperação de informações [15].

É possível que alguns estudos demonstrem funções monotónicas, lineares ou assintóticas, porém podem refletir a falha de não demonstrarem ou não abrangerem uma amplitude suficiente de intervalos. *Glenberg e Lehmann*, em 1980, conseguiram demonstrar que a relação entre o IIE e o IR mantinha-se mesmo quando estes eram medidos em dias ou semanas [20]. Este facto, foi mais tarde comprovado, de forma mais robusta, pelos estudos de *Cepeda* em 2008 e *Cepeda* em 2009[15][21]. Relativamente ao estudo realizado em 2008, utilizou-se itens triviais como material de estudo e usou-se o procedimento do teste/reestudo (teste com *feedback*). Os participantes estudaram os itens numa sessão em que apenas havia uma resposta correta para cada item apresentado. Após um IIE, que variava entre 0 e 105 dias, eles recebiam 2 testes/reestudo com o conjunto de itens anteriormente estudados. No final, recebiam um teste de recuperação para avaliar o desempenho de cada participante, após um IR que variou entre 7 e 350 dias.

*Glenberg e Cepeda*, mesmo utilizando métodos, procedimentos e intervalos diferentes, os resultados de *Cepeda* são muito próximos dos de *Glenberg*. Isto é, IIE ótimo torna-se mais longo à medida que o IR aumenta [13][22].

Todas estas descobertas foram realizadas em grande número em estudos de laboratório, contudo, faltava um estudo em grande escala aplicado à vida real, de forma a perceber se as ideias-chave se mantinham intactas. Desta forma, um estudo realizado em 2019, investigou esta questão, analisando uma base de dados longitudinal de cerca de 10514 funcionários de várias empresas, colhidas no contexto da formação local no trabalho. Os dados foram colhidos através de um sistema de gestão de aprendizagem - LMN, *learning management system* - que fornece conteúdo de treino a funcionários, dividido em várias sessões de aprendizagem. Uma característica importante

incorporada neste sistema é a capacidade de fazer variar apresentações de determinada informação, com intervalos de tempo variáveis e o número de itens apresentados em cada sessão, resultando em sessões de aprendizagem, naturalmente espaçadas no tempo de acordo com os horários de cada funcionário [23].

A primeira sessão corresponde ao primeiro contacto com o material de aprendizagem (material esse entregue a cada funcionário pela LMS), a segunda sessão corresponde ao segundo contacto com o material, momento de re-aprendizagem. Este intervalo de tempo entre as duas sessões é chamado, portanto, de intervalo interestudo -IIE. A terceira sessão corresponde ao momento de realização de um teste final. Assim, o intervalo entre a segunda sessão e a terceira, foi definida como intervalo de retenção - IR.

Segundo o presente estudo, chegou-se à conclusão que existe uma interação significativa entre o IIE e o IR. Estas conclusões estão de acordo com as descobertas feitas em laboratório. Demonstrou-se que a quantidade ideal de espaçamento entre uma sessão de aprendizagem inicial e uma sessão de re-aprendizagem varia, dependendo do quão longo é o IR. Mais especificamente, a probabilidade de reter informações na memória por um longo período de tempo - por exemplo, um mês ou mais - é maior se o intervalo IIE também for mais longo - por exemplo, 11 dias ou mais [21][23].

Em contrapartida, pequenos intervalos de espaçamento - por exemplo, 1 dia - é mais benéfico para um IR mais curto - por exemplo, 1 semana.

Assim, o intervalo ótimo de espaçamento entre uma sessão inicial de aprendizagem e uma sessão de re-aprendizagem, aumenta à medida que o IR também aumenta [23].

Portanto, o IIE deve ser aproximadamente 10%-20% do IR [15] [21] [23].

#### **4.1.3. Espaçamento igual ou sucessivamente maior**

Quando existem mais de duas sessões de estudo, os intervalos podem ser igualmente espaçados ou sucessivamente maiores – em expansão. Segundo os dados existentes, existem alguns estudos que apontam vantagens aos intervalos em expansão, enquanto outros mostram vantagens dos intervalos igualmente espaçados [12]. Portanto, mais estudos terão de ser realizados para esclarecer esta questão. De uma forma ou de outra, espaçar as sessões de estudo, trará sempre mais benefícios do que não o fazer – estudo massivo.

#### **4.1.4. Tipo de materiais**

O efeito do estudo distribuído tem vindo a ser observados em diferentes tipos de matérias a onde se pode incluir definições [24], pares rosto-nome de pessoas [25], traduções de palavras estrangeiras [26], fatos triviais [15], textos [14], palestras [27] e fotos [28]. O estudo espaçado também mostrou melhores desempenhos em vários domínios do conhecimento, tal como na biologia [29] e publicidade [30]. Por outro lado, se incluirmos capacidades técnicas, os domínios onde o efeito de

espaçamento foi demonstrado pelos seus benefícios, pode-se estender também para as matemáticas [31], história [32], música [33], cirurgia [34], entre outros.

#### **4.1.5. Características do aluno/ Abrangência**

A grande maioria dos estudos da prática distribuída é sobre estudantes, porém os efeitos também foram observados noutros tipos de população e faixas etárias. De forma genérica, as crianças de todas as idades beneficiam do estudo distribuído. Por exemplo, ao aprender imagens, as crianças desde a idade pré-escolar conseguem reconhecer e recordar mais itens estudados após intervalos mais longos entre sessões de aprendizagem, do que após curtos intervalos [35] [36]. Da mesma forma, crianças de 3 anos são mais capazes de classificar novos elementos de uma categoria, se a categoria foi inicialmente aprendida através de estudos espaçados, em vez de estudos massivos [37]. Até mesmo crianças de 2 anos mostraram benefícios com o estudo espaçado, de modo que aumentaram a sua posterior capacidade de produzir palavras anteriormente estudadas [38]. Também em crianças com deficiência específica na linguagem foi possível identificar benefícios [39].

Em relação ao outro extremo de idades, os adultos mais velhos, relativamente à aprendizagem de associação de pares de palavras, beneficiam igualmente do estudo distribuído tanto quanto os adultos jovens [40].

Também em certas populações clínicas, foi possível observar benefícios do estudo espaçado, incluindo indivíduos com esclerose múltipla [41], amnésia [42] e com lesões cerebrais traumáticas [43].

De forma geral, alunos de diferentes idades beneficiam do estudo distribuído, todavia existem ainda algumas perguntas em aberto. Desde logo, saber de que forma o conhecimento prévio ou a motivação individual podem influenciar os benefícios ou a magnitude do efeito do estudo distribuído.

#### **4.2. Efeito de Teste**

O efeito direto de testar é baseado em pesquisas que mostram que quando os alunos são testados numa certa matéria, conseguem recordar melhor do que quando não são testados.

A nível educacional, os testes são primariamente uma ferramenta utilizada pelos docentes para avaliar o conhecimento dos alunos e atribuir-lhes classificações. Contudo, existem cada vez mais estudos que indicam os testes como um instrumento que influencia diretamente a aprendizagem, promovendo uma melhor retenção de informação. Este efeito é conhecido como efeito de teste ou, do inglês, “*testing effect*” [44].

Estudos de laboratório de psicologia cognitiva têm demonstrado que, alunos que realizam testes com frequência e espaçados no tempo, após uma primeira sessão de estudo, apresentam

desempenhos de retenção de informação superiores em relação a estudantes, que reestudam a matéria no mesmo número de sessões de estudo [45].

#### **4.2.1. Efeito de teste versus reestudo**

Em duas experiências, um grupo de alunos, estudaram passagens de prosa e fizeram 1 ou 3 testes de recuperação imediatos, sem *feedback*, enquanto outro grupo reestudou as passagens o mesmo número de vezes. Depois, ambos os grupos receberam um teste final depois de 5 minutos, 2 dias ou 1 semana.

Quando o teste final foi fornecido após 5 min, o grupo de alunos que reestudaram as passagens tiveram melhores resultados no teste final, do que os que fizeram testes repetidos. Todavia, nos testes realizados 2 dias ou 1 semana depois, o grupo de alunos que fizeram testes repetidos tiveram um desempenho, substancialmente melhor do que os que reestudaram, embora estes relatassem o aumento da sua confiança na capacidade de lembrar a matéria. Assim, os testes, para além de se mostrarem um bom instrumento de avaliação de competências e de conhecimento, é uma ótima ferramenta para melhorar a aprendizagem [46].

Num outro estudo, internos de especialidade participaram numa sessão de ensino interativo sobre dois temas: Status epiléticos e Miastenia gravis. Foram distribuídos de forma randomizada por dois grupos equilibrados, em que num dos grupos, fizeram testes sobre o status epiléticos e estudaram a Miastenia gravis (SE-T/MG-S) e, o outro grupo o oposto, isto é, fizeram teste sobre Miastenia gravis e estudaram o status epiléticos (MG-T/SE-S).

Os testes e as sessões de estudo ocorreram imediatamente após a sessão de ensino e, posteriormente em 2 momentos adicionais, espaçados por 2 semanas entre cada momento. Os internos receberam *feedback* após cada teste. Os testes consistiam em perguntas de resposta curta, e as folhas de revisão de estudo tinham informações idênticas às respostas das perguntas abertas. Passado 6 meses, realizaram um exame final sobre os dois temas.

Em suma, 19 internos no grupo (SE-T/MG-S) e 21 internos do grupo (MG-T/SE-S) completaram o estudo. Analisando os resultados dos dois grupos, os internos que foram submetidos a testes repetidos tiveram resultados finais, que foram, em média, 13% melhores aos resultados obtidos pelos que reestudaram a matéria (39% vs 26%), 6 meses após a sessão inicial [47].

Estudos de neuroimagem realizados em 2016, demonstraram o aumento da conectividade entre as áreas do cérebro responsáveis pela atenção e a memória em estudantes durante um teste, comparativamente ao reestudo da matéria, tal como maior ativação de áreas cerebrais onde se dá o processamento semântico e/ou processamento da interpretação [48].

## **4.2.2. Efeitos primários do efeito de teste**

### **4.2.2.1. Característica do aluno/ Abrangência**

Existem estudos em que foi possível evidenciar o efeito de teste em diferentes populações de alunos e de diferentes faixas etárias, incluindo crianças do pré-escolar [49] crianças do ensino primário [50] [51]; alunos do ensino secundário [52] [53], alunos universitários [46] [54]. Porém, também no outro extremo de idades, verificou-se o efeito de teste, tanto em adultos de meia-idade como em adultos mais velhos [55] [56].

### **4.2.2.2. Tipos de matérias**

O efeito de teste foi demonstrado em diferentes materiais de aprendizagem, entre eles, listas de palavras e associações entre palavras [57], matérias mais complexas e em contexto de sala de aula, como citações de textos em prosa e interpretação [46] [58]. Foi identificado em informação fatural, incluindo fatos triviais e questões de conhecimento geral [59], tal como em fatos relacionados com matéria de ciências, história e psicologia [32] [60], desenvolvimento do adulto, neuroanatomia e história da arte [61]. Embora muitos dos estudos efetuados recaiam em matérias teóricas, foi possível também identificar em questões práticas, como na aprendizagem da competência de ressuscitação [62] e saturação laparoscópica [63].

## **4.2.3. Efeitos secundários ao efeito de teste**

Para além dos efeitos decorrentes diretamente do efeito de teste, discutidos anteriormente, foram identificados outros benefícios associados à prática de testes, que não advém diretamente da recuperação do conhecimento.

Num estudo, um conjunto de alunos foi dividido em dois grupos: alunos de um grupo foram submetidos a pequenos testes após as aulas, ao passo que outro grupo não. Os primeiros relataram que prestaram mais atenção às aulas, uma vez que sabiam que iam ser avaliados sobre a matéria apresentada. Para além disso, conseguiram identificar os tópicos mais importantes da matéria e monitorizar a sua aprendizagem, percebendo quais as suas lacunas e os seus pontos fortes [64].

Estudos recentes demonstram que alunos que são sujeitos a um teste após a exposição ao material de aprendizagem, apresentam um estudo subsequente muito mais eficiente, comparativamente com alunos que não o fizeram. Assim, um aluno será mais eficiente a estudar a matéria que foi testado no teste [65].

Noutro estudo, indivíduos que foram testados para a recordação de uma lista de palavras apresentadas aleatoriamente, demonstraram ser capazes de não só recordar mais palavras da lista, num teste final realizado dias depois, mas também uma melhor capacidade em organizar essas palavras em categorias, comparado com indivíduos que passaram o mesmo período de tempo a estudar as mesmas listas de palavras [66].

O *feedback*, ou o conhecimento dos resultados das perguntas dos testes, é muito importante na aprendizagem dos alunos, uma vez que lhes permite identificar as áreas mais frágeis do seu

conhecimento, e assim, alocarem, de forma estratégica, mais tempo de estudo de forma a melhorarem [67].

Um estudo realizado em estudantes de medicina, concluiu que a prática de testes, tem um papel muito importante na auto-monotorização do conhecimento. Os alunos têm uma tendência para dedicar mais tempo a rever matérias que apresentam uma discrepância entre as suas expectativas de resposta a uma pergunta e o seu resultado. Assim, tendem a dedicar mais tempo a rever perguntas que eles responderam incorretamente, acreditando ter respondido corretamente, e quando respondem de forma correta a uma pergunta que sentiam não saber a resposta. Para além disso, existe também uma maior tendência a rever perguntas respondidas incorretamente do que de forma correta [68].

Existem outros estudos que mostram que se os testes forem realizados com mais frequência, os alunos tendem a dedicar mais tempo ao estudo e acabam por fazê-lo de forma mais frequente, ao invés de condensarem o estudo nos dias que antecedem os testes [46].

Outro aspeto a realçar, é o facto de haver uma forte evidencia de que, testes de recuperação de informação potenciam a aprendizagem numa sessão de reestudo subsequente. Quando são realizados testes, o benefício associado ao reestudar uma certa informação é melhorado. Tudo indica que isto acontece pelo aumento da organização mental da informação existente. As realizações de testes não só têm um impacto positivo na recuperação futura de conhecimento já adquirido, como também prepara os alunos para a aprendizagem de nova informação no futuro. Assim, a memória subsequente é melhorada pela prática de recuperação e pelo estudo espaçado, tendo em vista que a combinação destes dois efeitos é um intensificador potente da memória [69].

A transferência de conhecimento consiste na aplicação de conhecimento previamente adquirido a novos contextos e situações. Assim, este é o objetivo final da aprendizagem. Estudos recentes nesse sentido, evidenciaram que o efeito de teste promove a transferência da aprendizagem, através do contexto temporal e de diferentes domínios do conhecimento. Portanto, a realização de testes ao longo do processo de aprendizagem, produz transferência de aprendizagem superior, quando comparado com o reestudo [70] [71].

#### **4.2.4. Formas de potenciar o efeito de teste**

##### **4.2.4.1. Formato de perguntas dos testes**

De forma geral, as evidências apontam para o facto de que testes com perguntas de resposta escrita são melhores do que testes de reconhecimento - como perguntas de escolha múltipla.

Num estudo realizado em contexto de sala de aula, os participantes foram convidados a assistirem a uma sequência de três aulas em dias consecutivos. Após cada um dos dias, foram submetidos a três tipos de atividades diferentes: estudar um resumo da aula, fazer um teste de escolha múltipla e fazer um exame de resposta curta. As respostas corretas foram apresentadas para metade das perguntas do teste de escolha múltipla e de resposta curta. Posteriormente, passado um mês, realizaram um teste de resposta curta. Depois de analisar os resultados, conclui-se que, fazer um

teste de escolha múltipla ou de resposta aberta, aumenta, significativamente, a retenção da informação a longo prazo, em relação a não fazer nada (Figura 6). Contudo, fazer teste de resposta aberta, foi a atividade pós-aula que mostrou melhores resultados. Assim, a implementação em sala de aula de um teste, seja de escolha múltipla, seja de resposta curta, após as aulas, terá um impacto significativo na retenção da matéria das aulas a longo-prazo [61].

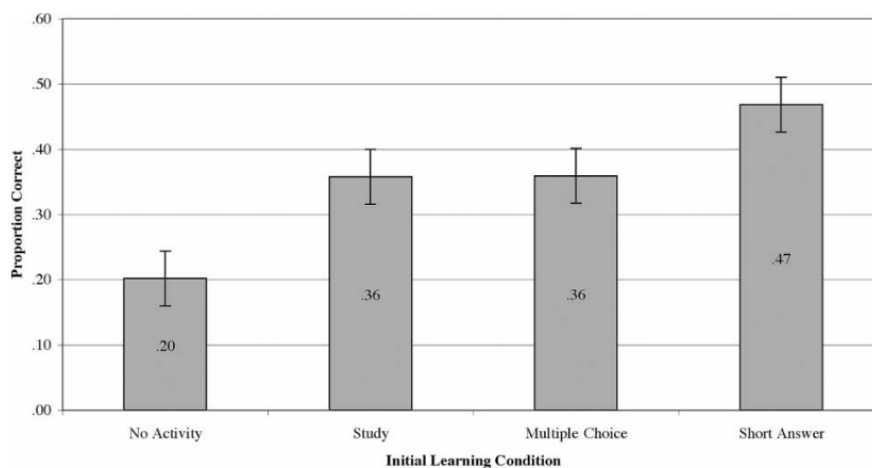


Figura 6 – Proporção de respostas corretas no teste de resposta curta final, em função da condição de aprendizagem pós-aula (as barras de erro apresentam um intervalo de confiança de 95%) [61].

Estas conclusões, estão de acordo com resultados de estudos anteriores de que, testes de resposta curta produzem retenções superiores da matéria das aulas, quando comparado com escolhas múltiplas. A condição de fazer testes de resposta curta, para além de produzir maiores retenções da matéria da aula, também concede maiores benefícios num teste subsequente, do que fazer um teste de reconhecimento inicial – testes de escolha múltipla [50] [72].

Igualmente tem sido evidenciado que, alunos que fazem testes frequentemente, como ferramenta de aprendizagem, apresentam níveis de ansiedade menores e aumento da regularidade de estudo [73].

#### **4.2.4.2. Feedback das perguntas**

Estudos realizados neste âmbito, levaram a concluir que o conteúdo do *feedback* é o aspeto mais importante em qualquer tipo de procedimento [74]. Um dos principais objetivos da investigação do *feedback* é determinar que tipo de informação contida na mensagem leva a uma maior efetividade na aprendizagem. A informação fornecida é decisiva para corrigir potenciais erros cometidos, assim como a manutenção e sedimentação das respostas corretas [75].

Num estudo tentou-se perceber de que forma o *feedback*, que fornece uma explicação da resposta correta, pode promover a transferência da aprendizagem para novas questões. Em duas experiências, os alunos estudaram e analisaram citações de prosa. De seguida, fizeram um teste

inicial de resposta curta sobre os conceitos do texto. Depois de cada questão respondida, recebiam *feedback* corretivo - isto é, se a resposta dada estava correta ou incorreta - ou *feedback* explicativo - informação adicional a justificar o porquê da resposta ser correta ou incorreta. Passados dois dias, os estudantes eram submetidos a um teste final constituído por perguntas repetidas do teste inicial e novas perguntas. De forma geral, o resultado demonstrou que o *feedback* corretivo e o *feedback* explicativo levavam a desempenhos semelhantes nas perguntas repetidas. Contudo, quando o teste final exigiu, através das novas perguntas, a transferência do conhecimento adquirido para um novo contexto, o *feedback* explicativo levou a um desempenho superior, em relação ao *feedback* corretivo. A informação adicional contida no *feedback* explicativo, leva a uma melhor e mais profunda compreensão de conceitos cruciais pelos alunos, aumentando assim a sua capacidade de os aplicar a novos contextos e a novas questões inferenciais [76].

O teste de escolha múltipla tem vindo a conquistar terreno nos diferentes sistemas educacionais, uma vez que é notório o seu alto valor pedagógico, fornecendo uma medida eficiente e eficaz da aprendizagem dos alunos. Tem ganho popularidade, muito graças ao avanço tecnológico, que permitiu tornar esta ferramenta de avaliação objetiva, rápida e fácil. Todavia, para além de todos os benefícios, os testes de escolha múltipla expõem os alunos a informação errada, através do uso de afirmações falsas nas várias opções de resposta. Durante a prática deste tipo de testes, os alunos podem adquirir conhecimentos errados. Uma das formas de maximizar os benefícios dos testes de escolha múltipla como ferramenta de aprendizagem e minimizar os efeitos negativos - aquisição de informação errada - passa por fornecer *feedback* após o teste. Ao fazê-lo, estamos a promover a aprendizagem ótima, auxiliando os estudantes a corrigirem os seus erros e a consolidar as respostas corretas [77].

Foi realizado um estudo com o intuito de perceber em que circunstâncias fornecer *feedback* pode ser utilizado para potenciar os benefícios dos testes de escolha múltipla, diminuindo os efeitos negativos, como a aquisição de informação errada. Nesse estudo, os alunos estudaram excertos de um texto e posteriormente receberam um teste de escolha múltipla com *feedback* imediato (logo após a resposta a cada questão), *feedback* atrasado (no final de responder a todas as questões) ou sem *feedback*. No geral, comparativamente à condição sem *feedback*, tanto o *feedback* imediato como o atrasado, aumentaram a proporção de respostas corretas e reduziram a proporção de invasões – opções de resposta erradas que interferiram com as respostas no teste subsequente - num teste final passado uns dias.

Foram avaliadas algumas variáveis neste estudo, como por exemplo: condição de estudo prévio (sem estudar, estudar uma vez, estudar duas vezes); o tipo de *feedback* fornecido (sem *feedback*, *feedback* imediato e *feedback* atrasado) e o número de alternativas de respostas. Primeiro, o efeito de teste é observado de forma significativa no teste final, isto é, os estudantes têm melhor desempenho em tópicos que foram testados no teste de escolha múltipla inicial, do que em tópicos que não foram testados, independentemente da condição de estudo prévio ou do *feedback* fornecido. Segundo, maior quantidade de estudo prévio leva a uma maior proporção de resposta

corretas no teste final, mas o número de alternativas de resposta não tem um impacto significativo nessa mesma proporção. Em contrapartida, menor quantidade de estudo prévio e um aumento do número de alternativas de resposta no teste de escolha múltipla inicial, resulta numa maior proporção de invasões. Terceiro, o *feedback* fornecido após um teste de escolha múltipla, aumentou o número de respostas corretas e diminuiu, drasticamente, a proporção de invasões. O *feedback* atrasado levou a uma maior proporção de respostas corretas comparado com o *feedback* imediato, mas ambos foram eficazes na redução da quantidade da informação errada adquirida [77].

Ambos os *feedbacks*, imediato e atrasado, mostraram grandes benefícios nas proporções de respostas corretas no teste final, comparativamente à condição sem *feedback*. Porém, o *feedback* atrasado tem maior vantagem em relação ao *feedback* imediato. Existem duas teorias diferentes, mas compatíveis, que podem explicar melhor essa vantagem. Em primeiro lugar, existe a teoria da interferência-preservação, que gira em torno da ideia de que o *feedback* imediato cria uma competição entre a resposta incorreta e a resposta correta apresentada pelo *feedback*. Desta forma o *feedback* atrasado acaba por ser mais eficaz, porque permite que respostas incorretas se dissipem, facilitando a aprendizagem da resposta correta [78] [79] [80]. A segunda teoria, reside na ideia de que o *feedback* atrasado leva a uma melhor recuperação da informação subsequente, porque permite fornecer uma apresentação espaçada adicional da matéria. Tal como já foi mencionado anteriormente, a superioridade do estudo distribuído em comparação com o estudo massivo está bem estabelecida na literatura [12]. De forma simples, considerando a resposta correta a uma questão do teste, como uma sessão de estudo, o *feedback* imediato representa a sessão de estudo massivo e o *feedback* atrasado, uma das sessões do estudo distribuído.

Por fim, a capacidade dos alunos em diferenciar entre respostas corretas e incorretas foi analisada através da correspondência absoluta entre as estimativas de confiança e a proporção de respostas corretas. Esta análise revelou que os alunos se apresentavam perfeitamente calibrados com as suas respostas sobre tópicos na condição de *feedback* atrasado, isto é, apresentavam alta confiança nas suas respostas, estando a proporção de respostas corretas de acordo com este nível, e apenas ligeiramente confiantes na condição de *feedback* imediato. Contudo, os alunos apresentavam-se altamente confiantes em relação às suas respostas na condição sem *feedback*, porém, a proporção de respostas corretas não estava de acordo com esta confiança. Ou seja, alunos que não recebem *feedback* tendem a apresentar uma falsa confiança sobre aquilo que realmente sabem e não sabem. Assim, o *feedback* auxilia os alunos a julgar e perceber melhor os seus pontos fortes e os seus pontos fracos [81].

É fulcral, portanto, que os professores forneçam *feedback* aos seus alunos, quando utilizam os testes de escolha múltipla como ferramenta de aprendizagem e de avaliação.

Não só nas respostas erradas o *feedback* se demonstrou importante, mas também nas respostas corretas. Um estudo realizado com esse propósito, mostrou que, quando as respostas corretas são dadas com baixa confiança, o *feedback* serve para corrigir o erro metacognitivo inicial, aumentando a retenção destas respostas corretas de baixa confiança. A metacognição é a consciencialização dos

alunos sobre os seus próprios conhecimentos e a sua capacidade de compreender e controlar as suas competências de aprender.

Em duas experiências, os alunos fizeram um teste de escolha múltipla sobre conhecimento geral de fatos. Fizeram um juízo sobre a sua confiança nas respostas dadas a cada pergunta. Foi fornecido *feedback* para metade das perguntas, e a retenção foi avaliada através de um teste final. Após a análise dos resultados, de forma geral, as duas experiências mostraram que fazer um teste de escolha múltipla inicial aumenta a retenção do conhecimento relativamente a não testar, sendo que fornecimento de *feedback* intensificou ainda mais o desempenho na retenção num teste final. Quando uma resposta incorreta é dada no teste de escolha múltipla inicial, é pouco provável que fosse corrigida, de forma espontânea, no teste final sem a apresentação de *feedback*. Logo, o *feedback* ajuda os alunos a corrigem erros de memória. Por outro lado, e mais importante, os resultados do estudo, demonstram que as respostas corretas beneficiam do *feedback*, sendo este efeito maior nas respostas corretas dadas com baixa confiança (Figura 7). Então, o *feedback* também ajuda os alunos a corrigirem o erro metacognitivo, que ocorre quando eles estão corretos em relação a uma pergunta num teste inicial, mas não têm confiança na sua resposta, resultando em maior retenção de respostas corretas dadas com baixa confiança. Por fim, também existe um benefício metacognitivo no teste final, melhorando a confiança das respostas dadas, de acordo com o real conhecimento dos alunos [59].

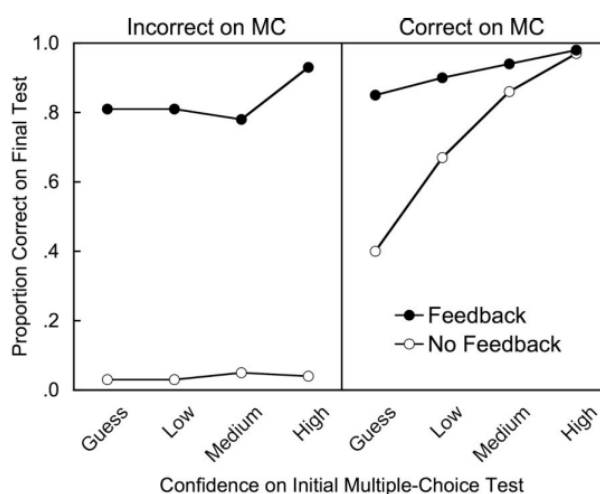


Figura 7 – Proporção de respostas corretas no teste final, em função da confiança de resposta e da condição de aprendizagem para respostas que estavam incorretas (lado esquerdo) e corretas (lado direito) no teste de escolha múltipla inicial (MC) [59].

#### 4.2.4.3. Tornar o efeito de teste explícito

Ao serem questionados acerca das estratégias de estudo utilizadas, apenas 11% dos alunos universitários utilizavam os auto-testes como instrumento de aprendizagem, e apenas 1% mencionou como estratégia preferida. Em contrapartida, 84% colocaram a releitura como uma das estratégias utilizadas, ao ponto que 55% como estratégia preferida. O estudo distribuído e a

utilização de testes como instrumento de aprendizagem e estudo, apresentam fortes evidência da sua grande utilidade na retenção de informação e conhecimento a longo prazo. Segundo este estudo, são as estratégias menos utilizadas [82].

Um das formas de incentivar os alunos a adotarem o auto-teste como ferramenta de estudo, pode passar por mostrar-lhes e fazê-los experimentar, de forma explícita, os benefícios e os efeitos do teste, uma vez que os alunos tendem a não o adotar de forma espontânea. Ou seja, os alunos devem perceber a melhoria da aprendizagem que os testes proporcionam, para mudarem as suas crenças metacognitivas e introduzirem cada vez mais o auto-teste às suas estratégias de estudo [83].

De forma a ilustrar o princípio mencionado, num estudo foi pedido aos alunos que lessem um texto utilizando a estratégia de Estudo-Estudo, e um texto usando a estratégia de Estudo-Teste. Passado uma semana, os alunos eram testados para ambos os textos com perguntas de resposta curta. Os resultados dos testes foram impressos e apresentados aos alunos, explicando a diferença entre os resultados obtidos após a estratégia Estudo-Estudo comparativamente à estratégia Estudo-Teste. Para além disso, fizeram uma revisão acerca das fortes evidências experimentais obtidas pelo efeito de teste. No final do semestre e após um questionário, 67% dos alunos avaliaram-se como um pouco mais ou muito mais propensos a usarem testes durante a leitura, enquanto 82% avaliaram-se como um pouco mais ou muito mais propensos a usar testes durante a leitura, comparativamente ao início do semestre.

Assim, esta estratégia poderia ser mais utilizada pelos professores no início de cada semestre, uma vez que, as evidências mostram que investir tempo e esforço a demonstrar e fazer os alunos experimentarem o efeito de teste, trará benefícios significativos para a aprendizagem [58].

#### **4.2.4.4. Cronograma de teste**

De forma a tornar o efeito de teste o mais eficiente e eficaz, as evidências apontam para que, numa fase inicial da aprendizagem, ocorra uma recuperação correta das informações, seguido por três a quatro testes subsequentes [22]. Por outro lado, os testes devem ser espaçados em intervalos sucessivamente mais longos, ao invés de intervalos igualmente espaçados, ao longo de dias ou semanas, para maximizar a retenção [85].

#### **4.2.4.5. Hipótese do esforço de recuperação**

O efeito de teste é robusto e muito estudado na literatura, porém ainda não existem certezas sobre o mecanismo pelo qual acontece.

Assim, um estudo colocou uma das hipóteses para esse mecanismo, a chamada hipótese do esforço de recuperação. Esta hipótese afirma que, recuperações de informação mais difíceis, mas bem-sucedidas, são melhores para a memória do que recuperações mais bem-sucedidas, mas mais fáceis. De forma a testar esta hipótese, estabeleceram-se condições em que a recuperação durante a

prática foi difícil. O IIE e o nível de critério (número de vezes que os itens foram pedidos a ser recuperados corretamente) foram manipulados para variar a dificuldade de recuperação.

De forma geral, os padrões dos resultados obtidos neste estudo confirmam as previsões da hipótese do esforço de recuperação. Assim, os resultados indicaram que, à medida que a dificuldade de recuperação durante a prática aumentava, o desempenho final do teste aumentou. IIEs mais longos comparados com mais curtos, leva a recuperações de informações mais difíceis, porém apresentam também desempenhos melhores no teste final. Além disso, à medida que o nível de critério aumenta, isto é, à medida que aumenta o número de vezes que os itens são recuperados corretamente, a recuperação é menos difícil e assim, observou-se retornos cada vez menores no desempenho no teste final -a partir da 6/7 vez que o item é recuperado.

Em suma, os resultados sugerem que as recuperações difíceis e bem-sucedidas são mais desejáveis do que as mais fáceis para a promoção da memória. Uma possibilidade que pode ajudar a justificar estes resultados é que, recuperações difíceis melhoram a variabilidade de codificação numa maior extensão do que as fáceis. A recuperação difícil pode envolver a ativação de informações mais relacionadas entre si, devido a uma pesquisa mais profunda e elaborada para a recuperação da informação desejada [84].



## 5. Exemplos da aplicabilidade na medicina

Todos os princípios e ideias-chaves desenvolvidas até agora, acerca do estudo distribuído e do efeito de teste, podem ser implementadas com ligeira facilidade a enumeras situações e fases do desenvolvimento dos estudantes de medicina e dos próprios médicos.

Estes princípios podem ser aplicados a diferentes níveis de conhecimento, seja ele teórico, prático ou até na modificação de comportamento e em diferentes contextos, seja na aprendizagem online, treino de simulação ou em sala de aula.

No contexto online, em termos de conhecimento teórico, temos o exemplo de um estudo em que se dividiram internos de urologia em dois coortes: o coorte 1 recebia três ciclos de estudo distribuído sobre a histopatologia da próstata-testículos (da semana 1 à 16) e três módulos condensados de ensino através da Web (WBT) sobre histopatologia da bexiga-rins (da semana 14 à 16). Os internos do coorte 2 o oposto, isto é, receberam três ciclos de estudo distribuído sobre a bexiga-rins (da semana 1 à 16) e três módulos condensados de WBT (da semana 14 à 16) (Figura 8). No estudo distribuído, diariamente os internos recebiam um email com um caso clínico com imagem histopatologica e questões sobre o diagnóstico. Os participantes recebiam *feedback* imediato depois de submeter a resposta. Os conteúdos dos emails eram iguais ao fornecido pelo WBT. A retenção a longo prazo foi avaliada durante as semanas 18 à 45. Após a análise dos resultados, concluiu-se que o estudo distribuído e o WBT tiveram em média um desempenho na retenção de longo prazo de 15,2% e 3,4%, respetivamente ( $p < 0,01$ ) (Figura 9). Embora o WBT apresente melhores resultados a curto prazo e com pico de desempenho mais precoce, o estudo espaçado tem, significativamente, maior retenção a longo prazo. Assim, concluiu-se que o estudo distribuído produz transferência das capacidades de diagnóstico e melhoram substancialmente a retenção a longo termo [86].

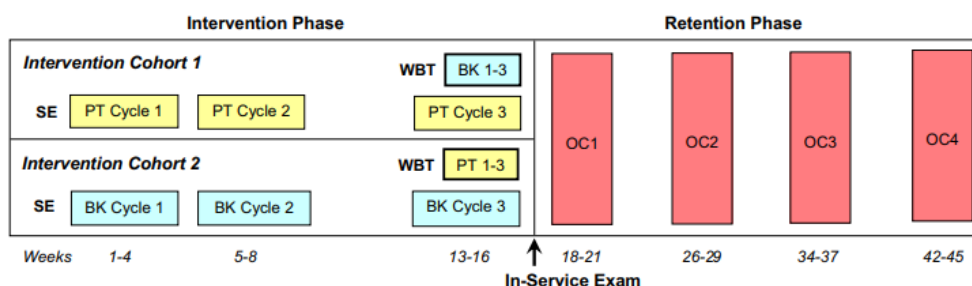


Figura 8 – Estrutura do ensaio clínico randomizado [86].

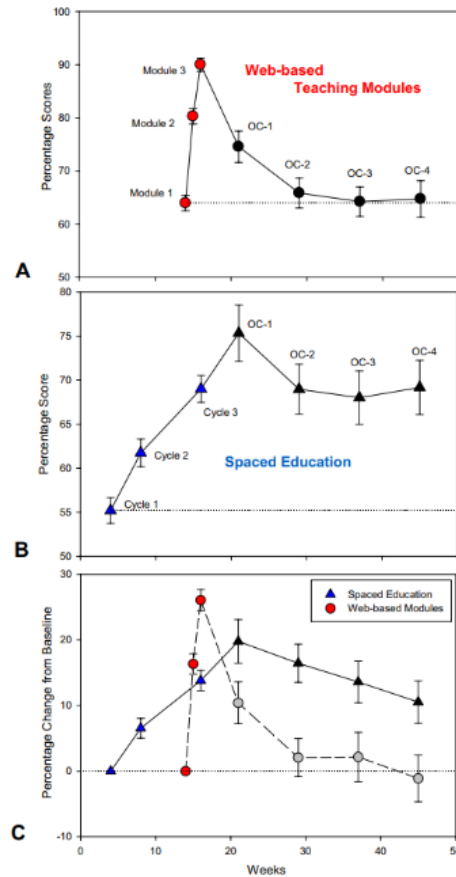


Figura 9 – Curvas de esquecimento para (A) ensino através da Web; (B) estudo espaçado; (C) ambos [86].

Outro exemplo, mas agora aplicado ao conhecimento prático, é um estudo onde se avaliou o impacto do estudo distribuído no desempenho da sutura laparoscópica. Para avaliar a efetividade do estudo distribuído, foi pedido aos participantes que executassem quatro nós de suturas num modelo de intestino, antes e após 3 horas de treino prático. Os vinte alunos foram, de forma randomizada, distribuídos por dois grupos. No grupo da aprendizagem distribuída, fizeram dois ciclos de treino durante 40 minutos, seguidos de uma pausa de 20 minutos onde aproveitavam para fazer uma caminhada. No grupo controlo, os participantes treinavam durante 3 horas seguidas sem interrupções. Todos os participantes eram estudantes de medicina e com pouca experiência na sutura laparoscópica. Foram avaliados o tempo total, a estabilidade do nó (avaliada por um tensímetro) a precisão da sutura, a qualidade do nó – (pontuação de *Muresan*) e o desempenho laparoscópico (lista de verificação de *Munz*). Também a motivação de cada participante foi avaliada através do Questionário da Motivação Atual.

Após a análise dos resultados, ambos os grupos melhoraram, significativamente, após o treino, em todos os aspetos avaliados. Contudo, os participantes que treinaram através do estudo distribuído, foram superiores em termos de desempenho de sutura, qualidade do nó e força da sutura (Figura 10, 11 e 12). Para além disso, é de realçar, que o estudo distribuído diminuiu, significativamente, a ansiedade e a impressão de desafio em comparação ao grupo controlo. Estes resultados podem

estar relacionados com o facto das pequenas pausas no estudo distribuído ajudassem a aumentar os níveis de concentração e motivação.

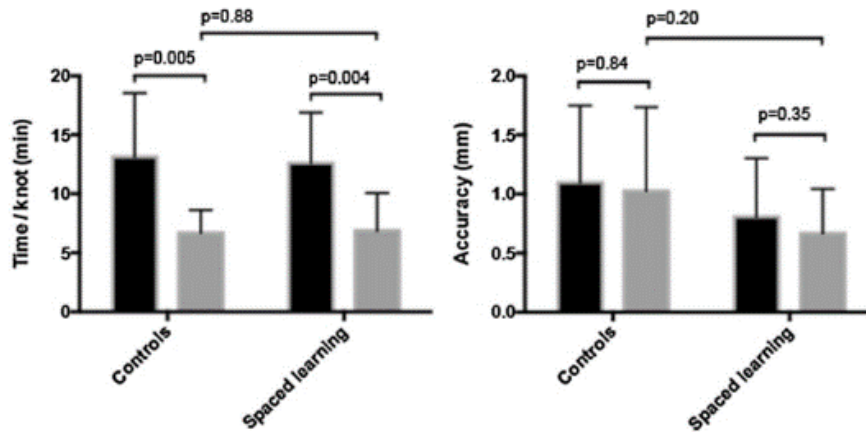


Figura 10 – Esquerda: todos os alunos aumentaram significativamente a velocidade de sutura e do nó – independentemente do método de treino. Direita: 3 horas de treino laparoscópico não afeta a precisão, significativamente [63].

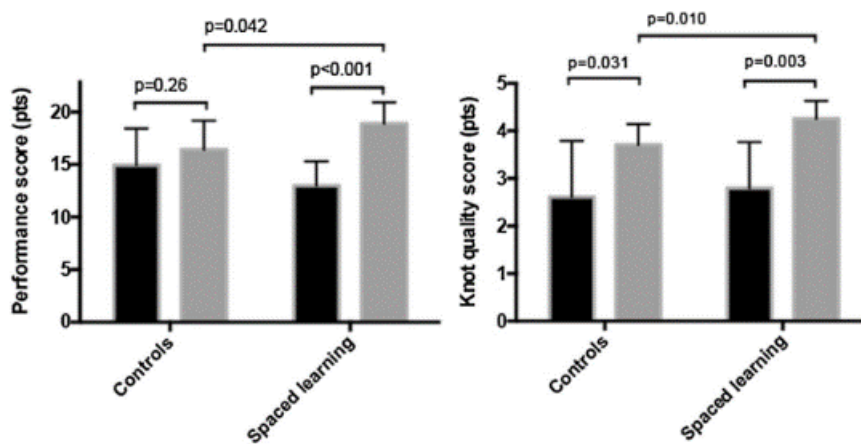


Figura 11 – Antes do treino, ambos os grupos tinham desempenhos quase semelhantes (esquerda) e pontuações de qualidade de nó (direita). A aprendizagem espaçada foi significativamente mais efetiva; apenas após três sessões espaçadas de 40 minutos (total de 120 minutos) comparativamente com 180 minutos de treino sem interrupções [63].

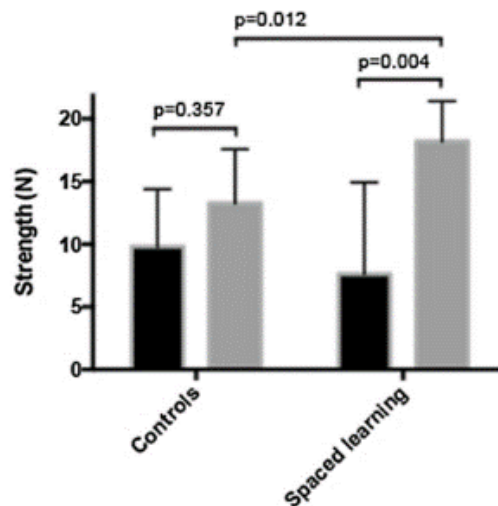


Figura 12 – No grupo do treino espaçado, o grau de força de sutura foi mais do que duplicada após 3 horas de treino, enquanto o grupo controlo não mostraram nenhuma melhoria significativa [63].

Em suma, o estudo distribuído apresenta evidência de ser muito útil, não só para o conhecimento prático, mas também para a aquisição de conhecimento motor complexo, como a suturação laparoscópica [63].

Outra aplicabilidade do estudo distribuído pode passar por ajudar a corrigir certos procedimentos, tal como a prescrição desnecessária de exames complementares diagnóstico.

Para tentar perceber de que forma o estudo distribuído poderia influenciar o número de rastreios desnecessários com o antigénio específico da próstata (PSA) prescritos pelos médicos, realizou-se o seguinte estudo. Os participantes foram randomizados em dois coortes: os médicos do coorte do estudo distribuído receberam ciclos igualmente espaçados de nove emails ao longo de 36 semanas - entre zero a dois emails por semana -, enquanto que o coorte de controlo não recebiam qualquer tipo de intervenção. Cada email apresentava um caso clínico, onde lhes era perguntado se o mesmo era apropriado para prescrever o rastreio, utilizando a PSA. Os participantes recebiam *feedback* imediato após o envio das respostas. A análise recaiu sobre o número e percentagem de rastreios PSA pedidos inadequadamente. Era considerado um pedido inadequado quando recaiam sobre doentes com >76 anos ou com <40 anos. Após a análise dos resultados, durante o período de intervenção (semana 1 à 36), os médicos do grupo do estudo distribuído, prescreveram menos rastreios inapropriados de PSA do que os do grupo controlo, (10,5% vs 14,2%,  $p=0,041$ ). Ao longo de um período de 72 semanas, após o período de intervenção, (semana 37 à 108), os médicos do grupo do estudo espaçado continuavam a pedir menos rastreios inapropriados de PSA do que os controlo (7,8% vs 13,1%,  $p=0,011$ ), representando uma redução relativa de 40% na prescrição de rastreios inapropriados (Figura 13). Portanto, o estudo distribuído melhora, de forma duradoura, o comportamento de rastreamento do cancro da próstata em médicos, mostrando-se ser uma metodologia promissora para melhorar os sistemas de saúde [87].

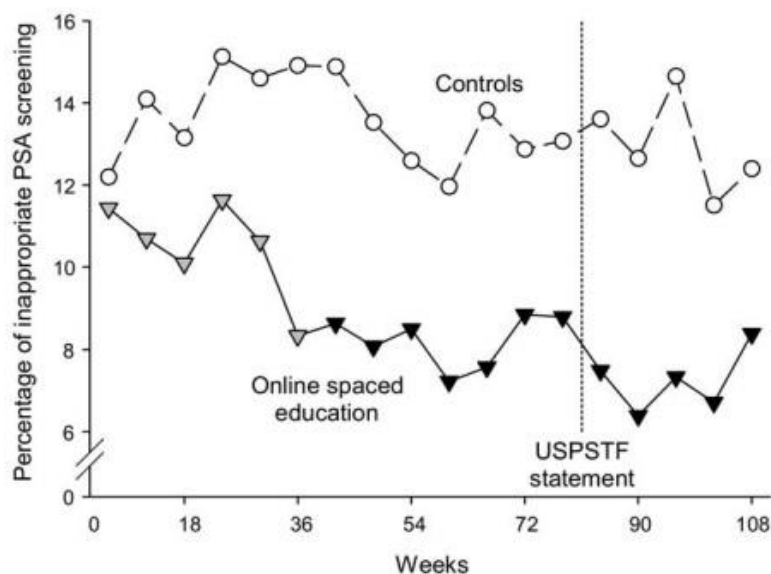


Figura 13 – Impacto da educação espaçada na triagem de PSA inadequada. As percentagens de rastreio de PSA inadequada durante e após a intervenção da educação espaçada são representadas pelos triângulos cinza e preto, respetivamente [87].

O estudo distribuído também tem sido aplicado após congressos ou palestras e tem apresentado resultados muito promissores. Tem se mostrado útil na consolidação das ideias-chaves discutidas nesses encontros, permitindo que os participantes os conseguissem reter durante maiores períodos de tempo [88] [89].

Resultados promissores também têm sido detetados em casos em que se adiciona o estudo distribuído às aulas tradicionais. Num estudo tentaram incorporar as duas vertentes. O estudo compreendia dois grupos de 21 estudantes do terceiro ano do curso de medicina dentária. Foram distribuídos, de forma randomizada, por um “grupo tradicional” e por um “grupo do estudo distribuído”. Ambos os grupos assistiam às aulas. No grupo do estudo espaçado a intervenção foi realizada de forma que esses alunos recebessem emails com um atraso de 14 dias para cada aula presencial (Figura 14). Os emails tinham perguntas de escolha múltipla sobre a matéria de aprendizagem das aulas. Após responderem às perguntas, recebiam *feedback* das respostas corretas e informação explicativa adicional. Todos os alunos, dos dois grupos, documentaram o tempo que dedicavam ao conteúdo de aprendizagem das diferentes aulas até à realização de um teste final no final do curso. Após a análise dos resultados, o grupo do estudo distribuído gastou, significativamente, mais tempo ( $216,2 \pm 123,9$  min) dedicado ao conteúdo da aprendizagem das aulas, comparado com o grupo tradicional ( $58,4 \pm 94,8$  min,  $p < 0,0005$ ). Para além disso, o grupo do estudo distribuído classificou o curso como sendo, significativamente, mais didático comparado com o grupo tradicional ( $p = 0,034$ ). Concluindo, ao adicionar o estudo distribuído às aulas presenciais tradicionais, incentiva os alunos a dedicarem mais tempo aos conteúdos das aulas [90].

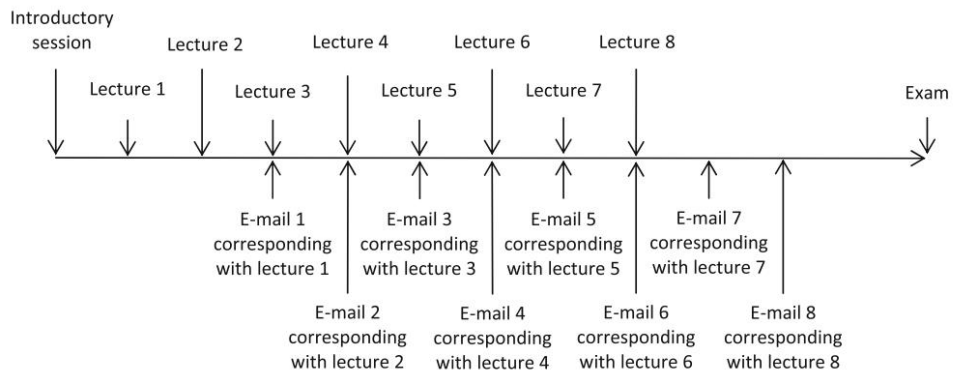


Figura 14 – Fluxograma do curso [90].

## 6. Problemas de implementação

Por muitas evidências da eficácia destes efeitos, existem sempre alguns problemas que podem complicar a sua implementação, seja no próprio sistema de educação, na forma de os professores darem as suas aulas, ou a forma como os alunos adotem estes métodos para a sua estratégia de estudo.

Assim, relativamente ao efeito do estudo distribuído, um dos problemas está relacionado com os livros didáticos que não incentivam o ensino distribuído. Isso acontece, uma vez que agrupam os temas por capítulos isolados, não fazendo revisões sistemáticas sobre temas que já foram ensinados, nos capítulos subsequentes. Da mesma forma, a maioria dos livros de matemática, por exemplo, no final de cada capítulo, têm um conjunto de exercícios só acerca desse tema, que nos capítulos subsequentes não voltarão a ser trabalhados [91].

Por outro lado, outros dos problemas reside na forma como os alunos tendem a estudar naturalmente. Num estudo [92], definiu-se um padrão típico da forma que os alunos estudam – “*procrastination scallop*”. Basicamente, o tempo gasto no estudo aumenta à medida que um teste se aproxima. Este padrão também foi identificado noutra estudo, onde voluntários concordaram em estudar numa sala de aula em observação, onde o tempo gasto a estudar era registado. Com testes realizados diariamente, os alunos tendiam a estudar por um período consistente de tempo entre cada sessão. Porém, quando o teste só era realizado passado 3 semanas, o tempo gasto com o estudo aumentou durante o intervalo, atingindo o pico imediatamente antes do teste. Assim, fazer testes com menor frequência leva a um estudo massivo no dia anterior ao teste. Ao passo que, testes mais frequentes, leva a um estudo mais distribuído pelo tempo. Pelo que parece, os alunos só se envolvem num estudo distribuído quando são forçados, não se sabe se isto se deve a restrições práticas ou porque não entendem os benefícios do estudo distribuído [93].

Em contrapartida, um estudo mostrou que os alunos conseguem ter uma maior consciência do seu nível geral de aprendizagem, após o estudo massivo do que após o estudo distribuído, mesmo quando experimentam os benefícios do distribuído [94]. No estudo massivo, como há a exposição a uma grande quantidade de informação, num curto período de tempo, os alunos ficam com a sensação de que aprenderam e sabem muito mais. De outro modo, no estudo distribuído, como a aprendizagem é feita de forma mais gradual e em períodos de tempo mais longos, os mesmos têm a sensação de estarem a aprender pouco e mais devagar. Outros estudos mais recentes, forneceram evidências de que os alunos não estão cientes dos benefícios de estudo distribuído, em contraste com estudo massivo [8] [11] [95].

Relativamente ao efeito de teste, como foi discutido anteriormente, o *feedback* é um fator de extrema importância. Assim, um dos problemas que existe, é o facto dos docentes, especialmente os universitários, muitas vezes não fornecerem *feedback* corretivo após os testes. Ou porque tentam tornar o acesso inoportuno (por exemplo, exigindo a marcação de um horário específico para verem os formulários dos testes), de forma a não perderem tempo de aula ou para poderem reutilizar as

perguntas em testes posteriores. Uma das sugestões poderia passar por os docentes fornecerem questionários de resposta curta, seguidos de *feedback* invés de fornecerem os resumos das aulas, ou os *powerpoints* [72].

Relativamente aos efeitos de teste, poderão existir algumas teorias que possam justificar a sua subutilização pelos alunos. Em grande parte, acredita-se que esse facto advém da existência de vieses cognitivos. Os alunos tendem a ser excessivamente confiantes na sua capacidade de recuperar informação numa situação posterior, em circunstâncias em que ainda é fácil processar essa informação (por exemplo, logo após estudar certo tema) – viés de fluência. Por outro lado, o viés de visão antecipada, acontece quando existem pistas durante a aprendizagem que não estão presentes no momento do teste [96] [97]. Já o viés da estabilidade, muito frequente, leva os alunos a possuírem a crença de que o grau de memória adquirida no momento da aprendizagem será mantido ao longo do tempo. Os alunos também podem ser vítimas da falácia da capacidade inata ou da crença, de que a diferença no desempenho entre os alunos, advém da aptidão ou da inteligência, e não da eficácia das estratégias de estudo e aprendizagem. Com isto, podem tornar-se menos motivados a continuar a aprender ou então, a perpetuarem estratégias de estudo menos eficazes [45] [98].

## 7. Possíveis soluções para implementação

Como possíveis soluções para ajudar na implementação destes efeitos na sala de aula e também nas estratégias de cada aluno, os docentes podem incorporar em cada aula uma breve revisão dos conceitos que foram apreendidos nas semanas anteriores, por exemplo. Os trabalhos de casa podem ser utilizados como forma de voltar a expor os alunos a informações e temas importantes que foram aprendidos anteriormente. Por exemplo, na disciplina de matemática, após uma aula, fornecer aos alunos, um conjunto de exercícios referentes não só à matéria da aula em questão, mas também a matérias de aulas anteriores. Ao intercalar os exercícios, contribuímos, inerentemente, para um estudo distribuído, para além que, forçamos os alunos a escolher a fórmula mais adequada para resolverem os diferentes problemas, algo que não acontece quando resolvem vários problemas referentes ao mesmo procedimento [99].

De outro modo, os docentes também podem optar por realizar exames ou questionários cumulativos, incentivando os alunos, por conta própria, a fazerem uma revisão dos capítulos anteriores [11] [54].

Referente ao efeito de teste, algumas das soluções para incentivar os alunos a testarem-se com regularidade, pode passar pelo uso de estratégias adequadas, tais como *flashcards*, que podem ser comercialmente preparadas pelos autores dos livros ou pelos próprios alunos. Cada vez mais, para além das questões de final de capítulo, os livros didáticos têm incluído, em versão digital, questões práticas de teste [8]. Para além disso, os estudantes devem também ser encorajados a terem um banco de perguntas onde possam praticar de forma intensiva, uma vez que os mesmos estão adaptados ao efeito de tempo, fornecendo *feedback* explicativo [45].

Por fim, existem evidências de que a aplicação em aula de um questionário semanal, apresenta melhorias significativas no desempenho no teste final de um curso, comparativamente a nenhum questionário. Num estudo equivalente, após um questionário com baixa peso na nota dos alunos, descobriram que o desempenho num teste final subsequente, melhorou 17%, comparado com quem não o fez [100]. Os mesmos efeitos foram encontrados em alunos mais jovens. Em todas as experiências, os alunos apresentaram melhorias de desempenho nos capítulos que foram testados anteriormente, comparado aos que não foram, e essa vantagem persistiu em exames surpresas realizados no final do semestre entre 1 a 3 meses depois [60].

Concluindo, os professores devem fazer mais testes e questionários nas suas aulas, e os alunos, adotarem cada vez mais autotestes durante o seu processo de aprendizagem.



## 8. Conclusões

Tanto o efeito de espaçamento como o efeito de teste, apresentam fortes evidências na literatura acerca da sua efetividade. Mesmo assim, continua a ser das estratégias de estudo menos utilizadas pelos estudantes em geral. A releitura de resumos, especialmente nos últimos dias que precedem o dia do teste, continua a ser muito disseminada pelos alunos.

Relativamente ao efeito de espaçamento, tende a ser próximo de ótimo quando o IIE é cerca de 10%-20% do IR. Por conseguinte, existe uma interação significativa entre eles, isto é, o IIE ótimo aumenta à medida que o IR aumenta. Por exemplo, para um IR de 1 semana, o IIE ótimo é de 1 dia, enquanto que um IR de um mês, o IIE ótimo é de 10-11 dias.

Alunos que por norma, não fazem testes durante o estudo, estão a perder benefícios que não conseguem de outra forma. Desde logo, fazer testes melhora, significativamente, o desempenho em testes de avaliação subsequentes e diminui os níveis de ansiedade e stress envolvidos. Trata-se de uma forma ativa e dinâmica de estudo, com níveis maiores de atenção envolvidos, comparado com a releitura, tornando o tempo de estudo altamente eficiente. Por outro lado, alunos que fazem testes com frequência, são melhores a monitorizar o seu nível de conhecimento, uma vez que conseguem identificar as suas lacunas e assim dedicar, de forma estratégica, mais tempo a melhorá-las. O efeito de teste promove uma maior capacidade de aplicar os conhecimentos adquiridos a novas situações - transferência da aprendizagem.

De forma a potenciar o efeito de teste, existem algumas variáveis a ter em conta. Preferir perguntas abertas a perguntas de escolha múltipla, numa fase inicial da aprendizagem. Utilizar *feedback* explicativo sempre que possível e de preferência atrasado, isto é, após a resposta a várias perguntas ao invés de ser imediato – após a resposta a cada pergunta. Optar por espaçar cada momento de teste, com um intervalo suficiente, de forma a que a recuperação da informação seja, desejavelmente, mais difícil do que fácil. Por fim, a abordagem mais custo-efetiva, passa por fazer uma recuperação bem sucedida, na fase inicial da aprendizagem, seguida de 3 ou 4 testes subsequentes, de preferência aumentando o período de tempo entre cada um deles.

No entanto, é necessário mais estudos que esclareçam o impacto de algumas condições, como é o caso do conhecimento prévio e os níveis de motivação, nos efeitos em causa.

Um exemplo prático: após a leitura de um capítulo sobre Tromboembolismo pulmonar, responder de forma breve a perguntas como: “ Qual o mecanismo fisiopatológico?”; “Quais as manifestações clínicas?”; “ Como se faz o diagnóstico”; “Qual o tratamento?” entre outras, com o livro fechado, sem qualquer tipo de ajuda. Passado 2 dias, fazer um bloco de 10-15 casos clínicos seguidos e corrigi-los apenas no final, lendo tanto as explicações da opção correta como das opções incorretas. Repetir o processo passado 1 semana e depois passado 1 mês e assim sucessivamente. Se porventura, responderem de forma muito fácil aos casos clínicos, devem aumentar os intervalos de tempo entre cada sessão, se muito difícil, encurta-los.

Esta pode ser uma abordagem útil tanto para os alunos dos primeiros anos de medicina, como para os alunos que se estão a preparar para a PNA (Prova Nacional de Acesso).

Com esta revisão bibliográfica, é desejável que se desmistifique os testes como sendo, única e exclusivamente, uma ferramenta de avaliação, temível pela maioria dos alunos mas, seja antes vista, como um instrumento importantíssimo de aprendizagem.

A junção do efeito de espaçamento com o efeito de teste, resulta numa enorme forma de potencializar as capacidades dos alunos em aprender e reter as informações por longos períodos de tempo, combatendo a curva de esquecimento.

## Referências bibliográficas

- [1] Squire LR. Memory systems of the brain: A brief history and current perspective. *Neurobiol Learn Mem* 2004;82:171–7. <https://doi.org/10.1016/j.nlm.2004.06.005>.
- [2] Squire LR, Zola-Morgan S. The medial temporal lobe memory system. *Science* (80- ) 1991;253:1380–6. <https://doi.org/10.1126/science.1896849>.
- [3] Cascella M, Al Khalili Y. Short Term Memory Impairment. *StatPearls* 2019.
- [4] Ebbinghaus H. Memory: A Contribution to Experimental Psychology. *Ann Neurosci* 2013;20. <https://doi.org/10.5214/ans.0972.7531.200408>.
- [5] Roesler R. Molecular mechanisms controlling protein synthesis in memory reconsolidation. *Neurobiol Learn Mem* 2017;142:30–40. <https://doi.org/10.1016/j.nlm.2017.04.015>.
- [6] MacLeod S, Reynolds MG, Lehmann H. The mitigating effect of repeated memory reactivations on forgetting. *Npj Sci Learn* 2018;3. <https://doi.org/10.1038/s41539-018-0025-x>.
- [7] Shail MS. Using Micro-learning on Mobile Applications to Increase Knowledge Retention and Work Performance: A Review of Literature. *Cureus* 2019;11. <https://doi.org/10.7759/cureus.5307>.
- [8] Dunlosky J, Rawson KA, Marsh EJ, Nathan MJ, Willingham DT. Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychol Sci Public Interes Suppl* 2013;14:4–58. <https://doi.org/10.1177/1529100612453266>.
- [9] Toppino TC, Gerbier E. About practice. Repetition, spacing, and abstraction. vol. 60. 1st ed. Elsevier Inc.; 2014. <https://doi.org/10.1016/B978-0-12-800090-8.00004-4>.
- [10] Kang SHK. Spaced Repetition Promotes Efficient and Effective Learning: Policy Implications for Instruction. *Policy Insights from Behav Brain Sci* 2016;3:12–9. <https://doi.org/10.1177/2372732215624708>.
- [11] Carpenter SK, Cepeda NJ, Rohrer D, Kang SHK, Pashler H. Using Spacing to Enhance Diverse Forms of Learning: Review of Recent Research and Implications for Instruction. *Educ Psychol Rev* 2012;24:369–78. <https://doi.org/10.1007/s10648-012-9205-z>.
- [12] Cepeda NJ, Pashler H, Vul E, Wixted JT, Rohrer D. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychol Bull* 2006;132:354–80. <https://doi.org/10.1037/0033-2909.132.3.354>.

- [13] Pashler H, Rohrer D, Cepeda NJ, Carpenter SK. Enhancing learning and retarding forgetting: Choices and consequences. *Psychon Bull Rev* 2007;14:187–93. <https://doi.org/10.3758/BF03194050>.
- [14] Rawson KA, Kintsch W. Rereading effects depend on time of test. *J Educ Psychol* 2005;97:70–80. <https://doi.org/10.1037/0022-0663.97.1.70>.
- [15] Cepeda NJ, Vul E, Rohrer D, Wixted JT, Pashler H. Spacing effects in learning: A temporal ridgeline of optimal retention: Research article. *Psychol Sci* 2008;19:1095–102. <https://doi.org/10.1111/j.1467-9280.2008.02209.x>.
- [16] Crowder RG. Principles of Learning and Memory. *Princ Learn Mem* 2014. <https://doi.org/10.4324/9781315746944>.
- [17] Glenberg AM. Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Mem Cognit* 1979;7:95–112. <https://doi.org/10.3758/BF03197590>.
- [18] Toppino TC, Bloom LC. The Spacing Effect, Free Recall, and Two-Process Theory: A Closer Look. *J Exp Psychol Learn Mem Cogn* 2002;28:437–44. <https://doi.org/10.1037/0278-7393.28.3.437>.
- [19] Verhoeven PPJL, Rikers RMJP, Schmidt HG. The effects of prior knowledge on study-time allocation and free recall: Investigating the discrepancy reduction model. *J Psychol Interdiscip Appl* 2005;139:67–79. <https://doi.org/10.3200/JRLP.139.1.67-79>.
- [20] Glenberg AM, Lehmann TS. Spacing repetitions over 1 week. *Mem Cognit* 1980;8:528–38. <https://doi.org/10.3758/BF03213772>.
- [21] Cepeda NJ, Coburn N, Rohrer D, Wixted JT, Mozer MC, Pashler H. Optimizing distributed practice theoretical analysis and practical implications. *Exp Psychol* 2009;56:236–46. <https://doi.org/10.1027/1618-3169.56.4.236>.
- [22] Rawson KA, Dunlosky J. When Is Practice Testing Most Effective for Improving the Durability and Efficiency of Student Learning? *Educ Psychol Rev* 2012;24:419–35. <https://doi.org/10.1007/s10648-012-9203-1>.
- [23] Kim ASN, Wong-Kee-You AMB, Wiseheart M, Rosenbaum RS. The spacing effect stands up to big data. *Behav Res Methods* 2019;51:1485–97. <https://doi.org/10.3758/s13428-018-1184-7>.
- [24] Dempster FN. The Spacing Effect: A Case Study in the Failure to Apply the Results of Psychological Research. *Am Psychol* 1988;43:627–34. <https://doi.org/10.1037/0003-066X.43.8.627>.

- [25] Carpenter SK, DeLosh EL. Application of the testing and spacing effects to name learning. *Appl Cogn Psychol* 2005;19:619–36. <https://doi.org/10.1002/acp.1101>.
- [26] Bahrick HP, Hall LK. The importance of retrieval failures to long-term retention: A metacognitive explanation of the spacing effect. *J Mem Lang* 2005;52:566–77. <https://doi.org/10.1016/j.jml.2005.01.012>.
- [27] Glover JA, Corkill AJ. Influence of Paraphrased Repetitions on the Spacing Effect. *J Educ Psychol* 1987;79:198–9. <https://doi.org/10.1037/0022-0663.79.2.198>.
- [28] Hintzman DL, Rogers MK. Spacing effects in picture memory. *Mem Cognit* 1973;1:430–4. <https://doi.org/10.3758/BF03208903>.
- [29] Reynolds JH, Glaser R. Effects of repetition and spaced review upon retention of a complex learning task. *J Educ Psychol* 1964;55:297–308. <https://doi.org/10.1037/h0040734>.
- [30] Appleton-Knapp SL, Bjork RA, Wickens TD. Examining the spacing effect in advertising: Encoding variability, retrieval processes, and their interaction. *J Consum Res* 2005;32:266–76. <https://doi.org/10.1086/432236>.
- [31] Rickard TC, Lau JSH, Pashler H. Spacing and the transition from calculation to retrieval. *Psychon Bull Rev* 2008;15:656–61. <https://doi.org/10.3758/PBR.15.3.656>.
- [32] Carpenter SK, Pashler H, Cepeda NJ. Using tests to enhance 8th grade students' retention of U.S. history facts. *Appl Cogn Psychol* 2009;23:760–71. <https://doi.org/10.1002/acp.1507>.
- [33] Simmons AL. Distributed practice and procedural memory consolidation in musicians' skill learning. *J Res Music Educ* 2012;59:357–68. <https://doi.org/10.1177/0022429411424798>.
- [34] Moulton CAE, Dubrowski A, MacRae H, Graham B, Grober E, Reznick R. Teaching surgical skills: What kind of practice makes perfect? A randomized, controlled trial. *Ann Surg* 2006;244:400–7. <https://doi.org/10.1097/01.sla.0000234808.85789.6a>.
- [35] Toppino TC. The spacing effect in young children's free recall: Support for automatic-process explanations. *Mem Cognit* 1991;19:159–67. <https://doi.org/10.3758/BF03197112>.
- [36] Toppino TC, Kasserman JE, Mracek WA. The effect of spacing repetitions on the recognition memory of young children and adults. *J Exp Child Psychol* 1991;51:123–38. [https://doi.org/10.1016/0022-0965\(91\)90079-8](https://doi.org/10.1016/0022-0965(91)90079-8).
- [37] Vlach HA, Sandhofer CM, Kornell N. The spacing effect in children's memory and category induction. *Cognition* 2008;109:163–7. <https://doi.org/10.1016/j.cognition.2008.07.013>.

- [38] Childers JB, Tomasello M. Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Dev Psychol* 2002;38:967–78. <https://doi.org/10.1037/0012-1649.38.6.967>.
- [39] Riches NG, Tomassello M, Conti-Ramsden G. Verb learning in children with SLI: Frequency and spacing effects. *J Speech, Lang Hear Res* 2005;48:1397–411. [https://doi.org/10.1044/1092-4388\(2005/097\)](https://doi.org/10.1044/1092-4388(2005/097)).
- [40] Balota DA, Duchek JM, Paullin R. Age-related differences in the impact of spacing, lag, and retention interval. *Psychol Aging* 1989;4:3–9. <https://doi.org/10.1037//0882-7974.4.1.3>.
- [41] Goverover Y, Hillary FG, Chiaravalloti N, Arango-Lasprilla JC, Deluca J. A functional application of the spacing effect to improve learning and memory in persons with multiple sclerosis. *J Clin Exp Neuropsychol* 2009;31:513–22. <https://doi.org/10.1080/13803390802287042>.
- [42] Cermak LS, Verfaellie M, Lanzoni S, Mather M, Chase KA. Effect of spaced repetitions on amnesia patients' recall and recognition performance. *Neuropsychology* 1996;10:219–27. <https://doi.org/10.1037/0894-4105.10.2.219>.
- [43] Goverover Y, Hillary F, Chiaravalloti ND, Arango-Lasprilla JC, DeLuca J. Application of the spacing effect to improve learning and memory.pdf. *J Clin Exp Neuropsychol* 2009;31:513–22.
- [44] Larsen DP, Butler AC, Roediger HL. Test-enhanced learning in medical education. *Med Educ* 2008;42:959–66. <https://doi.org/10.1111/j.1365-2923.2008.03124.x>.
- [45] Binks S. Testing enhances learning: A review of the literature. *J Prof Nurs* 2018;34:205–10. <https://doi.org/10.1016/j.profnurs.2017.08.008>.
- [46] Roediger HL, Karpicke JD. Test-enhanced learning: Taking memory tests improves long-term retention. *Psychol Sci* 2006;17:249–55. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>.
- [47] Larsen DP, Butler AC, Roediger HL. Repeated testing improves long-term retention relative to repeated study: A randomised controlled trial. *Med Educ* 2009;43:1174–81. <https://doi.org/10.1111/j.1365-2923.2009.03518.x>.
- [48] van den Broek G, Takashima A, Wiklund-Hörnqvist C, Karlsson Wirebring L, Segers E, Verhoeven L, et al. Neurocognitive mechanisms of the “testing effect”: A review. *Trends Neurosci Educ* 2016;5:52–66. <https://doi.org/10.1016/j.tine.2016.05.001>.
- [49] Fritz CO, Morris PE, Nolan D, Singleton J. Expanding retrieval practice: An effective aid to preschool children's learning. *Q J Exp Psychol* 2007;60:991–1004. <https://doi.org/10.1080/17470210600823595>.

- [50] Jaeger A, Eisenkraemer RE, Stein LM. Test-enhanced learning in third-grade children. *Educ Psychol* 2015;35:513–21. <https://doi.org/10.1080/01443410.2014.963030>.
- [51] McDaniel MA, Thomas RC, Agarwal PK, Mcdermott KB, Roediger HL. Quizzing in Middle-School Science: Successful Transfer Performance on Classroom Exams. *Appl Cogn Psychol* 2013;27:360–72. <https://doi.org/10.1002/acp.2914>.
- [52] Marsh EJ, Agarwal PK, Roediger HL. Memorial Consequences of Answering SAT II Questions. *J Exp Psychol Appl* 2009;15:1–11. <https://doi.org/10.1037/a0014721>.
- [53] Duchastel PC, Nungester RJ. Testing effects measured with alternate test forms. *J Educ Res* 1982;75:309–13. <https://doi.org/10.1080/00220671.1982.10885400>.
- [54] Carpenter SK, Pashler H. Testing beyond words: Using tests to enhance visuospatial map learning. *Psychon Bull Rev* 2007;14:474–8. <https://doi.org/10.3758/BF03194092>.
- [55] Balota DA, Duchek JM, Sergent-Marshall SD, Roediger HL. Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer’s disease. *Psychol Aging* 2006;21:19–31. <https://doi.org/10.1037/0882-7974.21.1.19>.
- [56] Maddox GB, Balota DA, Coane JH, Duchek JM. The Role of Forgetting Rate in Producing a Benefit of Expanded Over Equal Spaced Retrieval in Young and Older Adults. *Psychol Aging* 2011;26:661–70. <https://doi.org/10.1037/a0022942>.
- [57] Karpicke JD, Grimaldi PJ. Retrieval-Based Learning: A Perspective for Enhancing Meaningful Learning. *Educ Psychol Rev* 2012;24:401–18. <https://doi.org/10.1007/s10648-012-9202-2>.
- [58] Einstein GO, Mullet HG, Harrison TL. The Testing Effect: Illustrating a Fundamental Concept and Changing Study Strategies. *Teach Psychol* 2012;39:190–3. <https://doi.org/10.1177/0098628312450432>.
- [59] Butler AC, Karpicke JD, Roediger HL. Correcting a Metacognitive Error: Feedback Increases Retention of Low-Confidence Correct Responses. *J Exp Psychol Learn Mem Cogn* 2008;34:918–28. <https://doi.org/10.1037/0278-7393.34.4.918>.
- [60] McDaniel MA, Agarwal PK, Huelser BJ, Mcdermott KB, Roediger HL. Test-Enhanced Learning in a Middle School Science Classroom: The Effects of Quiz Frequency and Placement. *J Educ Psychol* 2011;103:399–414. <https://doi.org/10.1037/a0021782>.
- [61] Butler AC, Roediger HL. Testing improves long-term retention in a simulated classroom setting. *Eur J Cogn Psychol* 2007;19:514–27. <https://doi.org/10.1080/09541440701326097>.

- [62] Kromann CB, Jensen ML, Ringsted C. The effect of testing on skills learning. *Med Educ* 2009;43:21–7. <https://doi.org/10.1111/j.1365-2923.2008.03245.x>.
- [63] Boettcher M, Boettcher J, Mietzsch S, Krebs T, Bergholz R, Reinshagen K. The spaced learning concept significantly improves training for laparoscopic suturing: a pilot randomized controlled study. *Surg Endosc* 2018;32:154–9. <https://doi.org/10.1007/s00464-017-5650-6>.
- [64] Lyle KB, Crawford NA. Retrieving Essential Material at the End of Lectures Improves Performance on Statistics Exams. *Teach Psychol* 2011;38:94–7. <https://doi.org/10.1177/0098628311401587>.
- [65] Soderstrom NC, Bjork RA. Testing facilitates the regulation of subsequent study time. *J Mem Lang* 2014;73:99–115. <https://doi.org/10.1016/j.jml.2014.03.003>.
- [66] Zaromb FM, Roediger HL. The testing effect in free recall is associated with enhanced organizational processes. *Mem Cogn* 2010;38:995–1008. <https://doi.org/10.3758/MC.38.8.995>.
- [67] Son LK, Kornell N. Research on the Allocation of Study Time. *Handb Metamemory Mem* 2015. <https://doi.org/10.4324/9780203805503.ch17>.
- [68] Agrawal S, Norman GR, Eva KW. Influences on medical students' self-regulated learning after test completion. *Med Educ* 2012;46:326–35. <https://doi.org/10.1111/j.1365-2923.2011.04150.x>.
- [69] Arnold KM, McDermott KB. Free recall enhances subsequent learning. *Psychon Bull Rev* 2013;20:507–13. <https://doi.org/10.3758/s13423-012-0370-3>.
- [70] Butler AC. Repeated testing produces superior transfer of learning relative to repeated studying. *J Exp Psychol Learn Mem Cogn* 2010;36:1118–33. <https://doi.org/10.1037/a0019902>.
- [71] Carpenter SK. Testing Enhances the Transfer of Learning. *Curr Dir Psychol Sci* 2012;21:279–83. <https://doi.org/10.1177/0963721412452728>.
- [72] Kang SHK, McDermott KB, Roediger HL. Test format and corrective feedback modify the effect of testing on long-term retention. *Eur J Cogn Psychol* 2007;19:528–58. <https://doi.org/10.1080/09541440601056620>.
- [73] Leeming FC. The Exam-A-Day Procedure Improves Performance in Psychology Classes. *Teach Psychol* 2002;29:210–2. [https://doi.org/10.1207/S15328023TOP2903\\_06](https://doi.org/10.1207/S15328023TOP2903_06).
- [74] Shute VJ. Focus on formative feedback. *Rev Educ Res* 2008;78:153–89. <https://doi.org/10.3102/0034654307313795>.

- [75] Rohrer D, Pashler H, Cepeda NJ, Wixted JT. When does Feedback Facilitate Learning of Words? *J Exp Psychol Learn Mem Cogn* 2005;31:3–8.
- [76] Butler AC, Godbole N, Marsh EJ. Explanation feedback is better than correct answer feedback for promoting transfer of learning. *J Educ Psychol* 2013;105:290–8. <https://doi.org/10.1037/a0031026>.
- [77] Butler AC, Marsh EJ, Goode MK, Roediger HL. When additional multiple-choice lures aid versus hinder later memory. *Appl Cogn Psychol* 2006;20:941–56. <https://doi.org/10.1002/acp.1239>.
- [78] Kulhavy RW, Stock WA. Feedback in written instruction: The place of response certitude. *Educ Psychol Rev* 1989;1:279–308. <https://doi.org/10.1007/BF01320096>.
- [79] Kulhavy RW, Anderson RC. Delay-retention effect with multiple-choice tests. *J Educ Psychol* 1972;63:505–12. <https://doi.org/10.1037/h0033243>.
- [80] Kulhavy RW. Feedback in Written Instruction. *Rev Educ Res* 1977;47:211–32. <https://doi.org/10.3102/00346543047002211>.
- [81] Butler AC, Roediger HL. Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Mem Cogn* 2008;36:604–16. <https://doi.org/10.3758/MC.36.3.604>.
- [82] D. KJ, C. BA, L. RH. Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory* 2009;17:471–9.
- [83] Tullis JG, Finley JR, Benjamin AS. Metacognition of the testing effect: Guiding learners to predict the benefits of retrieval. *Mem Cogn* 2013;41:429–42. <https://doi.org/10.3758/s13421-012-0274-5>.
- [84] Pyc MA, Rawson KA. Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *J Mem Lang* 2009;60:437–47. <https://doi.org/10.1016/j.jml.2009.01.004>.
- [85] Kang SHK, Lindsey R V., Mozer MC, Pashler H. Retrieval practice over the long term: Should spacing be expanding or equal-interval? *Psychon Bull Rev* 2014;21:1544–50. <https://doi.org/10.3758/s13423-014-0636-z>.
- [86] Kerfoot BP, Fu Y, Baker H, Connelly D, Ritchey ML, Genega EM. Online spaced education generates transfer and improves long-term retention of diagnostic skills: A randomized controlled trial. *J Am Coll Surg* 2010;211. <https://doi.org/10.1016/j.jamcollsurg.2010.04.023>.

- [87] B.P. K, E.V. L, G. S, D. G, P.R. C. Durable improvements in prostate cancer screening from online spaced education: A randomized controlled trial. *Am J Prev Med* 2010;39:472–8.
- [88] Shaw B. T, Long B. A, Chopra B. S, Kerfoot B. P. Impact on Clinical Behavior of Face-to-Face Continuing Medical Education Blended with Online Spaced Education: A Randomized Controlled Trial. *J Contin Educ Health Prof* 2011;31:103–8.
- [89] D.E. G, T. S, J. N, C. B, M. P, A. S, et al. Improving the impact of didactic resident training with online spaced education. *ANZ J Surg* 2013;83:477–80.
- [90] E. N, E. V, A. B, S. E, A. B, C. K. Spaced education activates students in a theoretical radiological science course: a pilot study. *BMC Med Educ* 2012;12:32.
- [91] Stigler JW, Ham M, Kim MS, Fuson KC. An Analysis of Addition and Subtraction Word Problems in American and Soviet Elementary Mathematics Textbooks. *Cogn Instr* 1986;3:153–71. [https://doi.org/10.1207/s1532690xcio303\\_1](https://doi.org/10.1207/s1532690xcio303_1).
- [92] Michael J. A Behavioral Perspective on College Teaching. *Behav Anal* 1991;14:229–39. <https://doi.org/10.1007/bf03392578>.
- [93] Mawhinney VT, Bostow DE, Laws DR, Blumenfeld GJ, Hopkins BL. A comparison of students studying-behavior produced by daily, weekly, and three-week testing schedules1. *J Appl Behav Anal* 1971;4:257–64. <https://doi.org/10.1901/jaba.1971.4-257>.
- [94] Kornell N, Bjork RA. Learning concepts and categories: Is spacing the “enemy of induction”? *Psychol Sci* 2008;19:585–92. <https://doi.org/10.1111/j.1467-9280.2008.02127.x>.
- [95] Wissman KT, Rawson KA, Pyc MA. How and when do students use flashcards? *Memory* 2012;20:568–79. <https://doi.org/10.1080/09658211.2012.687052>.
- [96] Koriat A, Bjork RA. Illusions of competence in monitoring one’s knowledge during study. *J Exp Psychol Learn Mem Cogn* 2005;31:187–94. <https://doi.org/10.1037/0278-7393.31.2.187>.
- [97] Tullis JG, Benjamin AS. On the effectiveness of self-paced learning. *J Mem Lang* 2011;64:109–18. <https://doi.org/10.1016/j.jml.2010.11.002>.
- [98] Bjork R. Institutional impediments to effective training. *Learn Rememb Believing Enhancing Individ Team Perform* 1994:295–306.
- [99] Rohrer D, Taylor K, Sholar B. Tests Enhance the Transfer of Learning. *J Exp Psychol Learn Mem Cogn* 2010;36:233–9. <https://doi.org/10.1037/a0017678>.

[100] Kang SHK, McDaniel MA, Pashler H. Effects of testing on learning of functions. *Psychon Bull Rev* 2011;18:998–1005. <https://doi.org/10.3758/s13423-011-0113-x>.

