Advances in Psychology 心理学进展, 2021, 11(3), 657-664
Published Online March 2021 in Hans. http://www.hanspub.org/journal/ap
https://doi.org/10.12677/ap.2021.113076

Hans汉斯

提取练习及其应用的研究综述

杨简绿

西南大学心理学部，重庆
Email: 820325675@qq.com

收稿日期：2021年2月7日；录用日期：2021年3月8日；发布日期：2021年3月17日

摘　要

提起学习策略，可能大部分人想到的还是重读、重学等比较低效的策略。而提取练习是指通过测试、自由回忆等方式从记忆中多次提取信息。相比于重学等学习策略，提取练习促进了对材料的学习，进一步提高记忆效果，增长记忆时间，这种效应被称为提取练习效应，也叫测试效应。本文围绕提取练习这一主题，主要从机制、局限性以及应用展望来梳理进行简要讨论。

关键词

提取练习，测试效应，记忆，综述

A Review of Retrieval Practice and Applications

Jianlv Yang

Faculty of Psychology, Southwest University, Chongqing
Email: 820325675@qq.com

Received: Feb. 7th, 2021; accepted: Mar. 8th, 2021; published: Mar. 17th, 2021

Abstract

When it comes to learning strategies, most people think of re-reading, re-study and other relatively inefficient strategies. The retrieval practice refers to multiple extraction of information from memory through testing, free recall and other methods. Compared with learning strategies such as re-study, retrieval practice promotes the learning of materials, further improves the memory performance, and increases the memory time. This effect is called the retrieval practice

文章引用: 杨简绿(2021). 提取练习及其应用的研究综述. *心理学进展, 11(3),* 657-664.
DOI: 10.12677/ap.2021.113076

effect or the testing effect. This paper briefly discusses the topic of retrieval practice, focusing on the mechanisms, limitations and application perspectives.

Keywords

Retrieval Practice, Testing Effect, Memory, Review

1. 提取练习

提取练习(retrieval practice, RP)指通过测试、自由回忆等方式从记忆中多次提取信息，因此提取练习效应(retrieval practice effect, RPE)也被称为测试效应(testing effect)，指提取练习促进了对学习材料的记忆，提高了再次回忆的能力(Carrier & Pashler, 1992; Carpenter & DeLosh, 2005; Roediger & Karpicke, 2006)。不同于大部分低效的学习记忆策略，如重复学习、划线(Dunlosky et al., 2013)，只是强调对学习材料进行多次复述，而提取练习更加高效，也更有适用性。也就是说，与之前提到的重复学习记忆策略相比，提取练习能够有更好的记忆效果与更长的保持时间。

前人研究与行为实验证明，提取练习对于学习和记忆表现的提高可以是直接的，也可以是间接的。实验证据表明，在学习过程中进行测试，有利于提高对于给定学习材料的随后记忆表现(Izawa, 1971; Karpicke et al., 2009; Roediger & Butler, 2011)，也就是说提取练习这一行为本身对学习记忆产生了直接影响，增强了对同一材料的后续学习，对记忆痕迹进行了强化，使该信息能够保持更好的记忆与更持久的时间(Knight et al., 2012)；除了能够提高对给定材料的直接记忆，提取练习还能间接促进对后续新材料的学习，因为提取练习促进了之后再次学习中对信息的编码活动，采用了更加有效的编码方式(Szpunar et al., 2008; Wissman et al., 2011)。

测试除了可以作为评价学习效果的一种手段，也可以作为学习记忆的一种比较高效的策略，同时也具有一定的普适性，首先，前人使用了多种材料、方式对提取练习进行了研究，包括面孔－名字配对(Carpenter & DeLosh, 2005)，外语词汇(Karpicke & Roediger, 2008)，标准单词配对(Carpenter, 2009)，一般事实知识(Carpenter, Pashler, & Cepeda, 2009)，SAT 美国学术能力评估测验问题(Marsh et al., 2009)，不同概念类别样本(Jacoby, Wahlheim, & Coane, 2010)，段落文本(Chan, 2010)，输出作答或回忆(Smith, Roediger & Karpicke, 2013)等，都发现了提取效应的存在。其次，提取练习不受年龄限制，前人研究发现，从学龄前儿童甚至到中老年人都有提取练习效应(Szpunar, Jing, & Schacter, 2014; Karpicke, Blunt, & Smith, 2016; Fritz, Morris, Nolan, & Singleton, 2007; Meyer & Logan, 2013)，除此之外，Rowland (2014)通过元分析发现，提取练习效应在各个年龄段的效应差异并不大。最后，不同能力水平都同样受到提取练习的增益，包括知识水平(周爱保，马小凤，李晶，崔丹，2013)，注意控制能力(Brewer & Unsworth, 2012)，记忆能力(Coyne, Borg, DeLuca, Glass, & Sumowski, 2015)，阅读能力(Karpicke, 2016)等。

由于提取练习的广泛适用性，不少研究者都开始扩展提取练习的应用领域，因此理解提取练习的机制以及局限性对于提取练习的应用是具有重要意义的，本文将就这几个方面对提取练习及其应用进行阐述。

2. 提取练习的相关机制假说

2.1. 努力提取假说

努力提取假说(Effort Retrieval Hypothesis)认为，相比于重学、识别等过程，提取这一过程需要更多的认知资源以及加工深度，可以提高对记忆的精加工以及增加信息的提取路径，因此记忆效果更好更持久(Bjork, 1975; Gardiner et al., 1973)。Bjork 认为，提取深度与编码加工深度相似，更加困难、深入的提取可以提高记忆效果。行为研究也提供了相应的证据，有研究发现，当第一次测试的线索与原始编码不匹配，但提取成功时，第二次测试出现的测试效应比第一次测试的线索和匹配时要大，研究者认为是由于记忆痕迹的提取路径增加或者可能是记忆痕迹本身的特征增加了(McDaniel & Masson, 1985)。另外有证据表明，需要更多时间来回想一般知识问题答案的人，在后面的测试中更容易回想出答案(Benjamin, Bjork, & Schwartz, 1998)。Pyc 和 Rawson (2009)通过增大测试间隔增加提取难度，发现随着提取难度加大，被试在后面的测试中表现更好，支持了努力提取假说。还有一项研究通过操纵编码时长和提取时长发现，不管在即时测试和延迟测试下，适当增加提取时长都发现了更高的记忆成绩(张锦坤、张俐娟，2020)。

Bjork 和 Bjork (1992)进一步扩展了理论，区分了储存强度(storage strength)和提取强度(retrieval strength)。储存强度反映了记忆痕迹或者学习的相对持久性，而提取强度反映了短时间内的记忆痕迹的可获得性(提取流畅性)，也就是反映了从记忆中的提取难度。Bjork 假设，储存强度与提取强度是呈负相关，具体来说，更容易的提取(提取强度低)并不会增加储存强度，但是需要更多努力的提取确实增加了储存强度，促进了更稳固更长久的记忆。

2.2. 中介效益假说

中介效益假说(Mediator Effectiveness Hypothesis)认为，测试手段通过在编码过程中产生并使用了更有效(与重学相比)的中介，因此更有可能在之后被提取和解码，从而增加了对目标信息的回忆并以此改善记忆表现，这里的中介指连接线索与目标的单词、短语或概念(Pyc & Rawson, 2010)。Pyc 和 Rawson 认为，有效的中介主要取决于两个关键因素：① 中介提取，指当线索信息出现时，中介是可以被回忆起来的；② 中介解码，就是中介能够从记忆中唤起目标信息。

Pyc 和 Rawson 通过一项行为实验来验证，在最初的单词学习后，他们将被试分为两组：① 测试－重学组，在每一个重学项目之后都有一个即时线索回忆测试；② 重学组，被试只是重新再次学习单词。并且要求所有被试在学习和重学阶段都要生成并报告帮助回忆的关键词中介。最后测试分为三组：① 线索组，只有线索；② 线索－中介组，线索和他们自己生成的关键词；③ 线索－中介提示组，线索和要求回忆中介词的提示。结果发现，在只有线索的情况下，测试-重学组记忆表现远好于重学组。在加入关键词后，极大地改善了重学组的记忆表现，而测试-重学组的表现没有明显差异，这表明在测试－重学组的被试自发地使用了有效的中介，因此提供关键词没有增益。最后，还发现了测试-重学组回忆出的关键词更多，正确率更高。

Pyc 和 Rawson 认为结果支持了中介效益假说，在练习中成功提取中介能够增加记忆强度，另外提取失败可能会促进将低效的中介转向更加有效的中介。他们也提到中介效益假说也许不是唯一机制解释，但是中介确实起到了比较重要的作用(Pyc & Rawson, 2010)。Miyatsu 和 McDaniel (2019)后来的一项研究也侧面证明了此假说，他们发现，将提取练习和关键词助记这两种策略结合能够极大地改善对外语词汇的记忆表现。

2.3. 精细提取假说

精细提取假说(Elaborative Retrieval Hypothesis)认为，与重新学习只涉及到对呈现信息的阅读相比，

提取行为则需要一个过程，在此过程中记忆的内容被激活，以努力寻找一些没有呈现的目标信息。而按照记忆的扩散激活理论(Collins & Quillian, 1972)，在提取过程中激活的信息可能会扩散到其他相关概念上，并最终激活一个具有多条路径通往正确目标的精细的语义网络，这种精细的网络结构有利于记忆保持，因为它提供了更多能够激活目标的信息(Carpenter, 2009)。

Carpenter 通过两个行为实验来验证精细提取假设，在实验中，被试完成对提示－目标词对的初始编码，包括强关联词对(吐司－面包)和弱关联词对(篮子－面包)，之后通过提示回忆测试(如：吐司－____)或重学(如：吐司－面包)进行练习，最后进行两次自由回忆测试，结果发现测试项目比重新学习项目的记忆表现得更好，虽然强关联促进了最初测试的记忆表现，但随着时间的推移，弱关联的项目保持得更好，以至于强关联优势在最后测试时被消除或逆转。而无论提示－目标关系的强弱，重新学习的项目在最后测试时的回忆率都是相似的。因此，研究者认为，与重学相比，提取练习过程更容易产生信息的激活形成精细语义网络。

3. 局限性

虽然提取练习有积极影响，如前文提到的增强记忆表现，有利于长时记忆保持，但是在某些情况下也可能产生消极影响。

3.1. 提取的干扰效应

虽然提取练习能够提高被测试的材料的记忆表现，但是有时候也会损害没有被测试的其他材料的提取。

不少实验研究都发现了提取诱发遗忘现象。在一项使用配对关联学习范式的实验中，研究者改变了测试项目的顺序，发现项目在较后测试时比较前测试时表现更差(Tulving & Arbuckle, 1966)；Roediger (1978)发现，在不同类别词表记忆后，给予了部分词表类别提示的被试虽然在提示类别下记忆表现比没有提示的被试更好，但是整体表现不如没有提示组，表明被提示的类别回忆干扰了其他类别的项目回忆；Anderson、Bjork 和 Bjork (1994)通过类别－目标词对进行实验发现，在练习阶段提取的项目记忆表现高于未提取的项目，但与提取项目有关却未被提取的项目比未提取的项目记忆表现更差，也就是出现了提取诱发遗忘现象(Anderson, Bjork, & Bjork, 1994)。

虽然有发现提取的干扰效应，但这种干扰并不是绝对存在的，比如在段落文本材料的测试中，提取部分文本可以促进其他未被提取的部分的记忆表现。也有学者提出，在输出－干扰范式中观察到的抑制作用通常不大，也就是说，提取虽然产生了输出干扰效应，但测试的整体效果还是积极的，虽然这种假设还未得到实验或者测试证实(Roediger & Karpicke, 2006)。

3.2. 消极暗示效应

Smith 等人(2013)提出提取练习效应的一个关键属性是取决于提取的成功，且有研究发现只有当被试对他们的答案自信程度较高时，他们才能从成功的提取练习中获益(Zhang, Chen, & Liu, 2019)。因此消极暗示效应是指如果错误地提取测试信息，很有可能在以后再次出现同样的错误。

在一项前人研究中，任务要求被试完成阅读过和未阅读过的段落的选择题测试，阅读过的段落的正确率高于未阅读过的段落，但是随着错误选项的增多，被试的表现也越来越差。后来研究者让被试进行了自由回忆测试，测试的题目包括做过选择题的题目和未做过选择题的题目，阅读和未阅读过的段落都有提取练习效应，但是也出现了消极暗示效应，因为对被测试项目的正确反应随着多项选择测试中错误选项数量的增多而下降(Marsh, Fazio, & Roediger, 2006)。

虽然提取错误会产生消极暗示效应，但相关研究证据表明如果在参加选择题测试后不久就得到反馈，则可以逆转了消极暗示效应(Butler & Roediger, 2006)。

4. 应用及展望

4.1. 学习应用及展望

提取练习的众多好处让许多研究人员提倡使用测试来促进学生的记忆保持(Bjork, 1988; Dempster, 1996; McDaniel, Roediger, & McDermott, 2007; Pashler, Rohrer, Cepeda, & Carpenter, 2007)。有研究支持了这一概念，在模拟课堂环境(Butler & Roediger, 2007)和实际课堂环境中通过测验学习课程，学习内容的记忆表现得到了改善，表现出可靠的测试效果(McDaniel, Anderson, Derbish, & Morrisette, 2007; Roediger & Karpicke, 2006; Carpenter, Pashler, & Cepeda, 2009)。

与重复学习相比，重复测试往往会减慢初始的学习速度，在刚开始的时候学习效果并不明显，但测试会促进更持久更稳固的记忆。可是由于学生经常使用当前处理信息的流畅性(提取强度)作为当前学习状况的证据，因此大部分学生可能会选择糟糕的学习策略，比如重读、划线等，也就是说学生一般会选择提高他们当前的流畅性而不是去考虑是否有利于长期学习(Jacoby, Bjork, & Kelly, 1994; Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013)。在许多情况下，能让初始学习快速增长的策略不利于记忆的长期保持；同样，更努力的策略往往会增强记忆的长期保持，比如提取练习。如果学生根据其当前的流畅性来监测和指导他们的学习，可能会受到能力错觉的影响，认为未来的表现大于实际表现(Bjork, 1999; Jacoby et al., 1994; Koriat & Bjork, 2005)；但如果对学生进行提取练习干预，告诉他们提取练习的好处以及方法，他们之后就会进行自我调节，自发使用提取练习改善学习效果(Ariel & Karpicke, 2018)。

但是提取练习是否是一种适合所有人或所有情况的学习策略还未有定论，最近一项研究发现，提取练习效应与人格特质、工作记忆之间容量之间的相关系数并不高(Bertilsson et al., 2020)。而之前有研究表明，相比于工作记忆容量高的学生，提取练习对于工作记忆容量较低的学生来说，可能是特别有效的策略(Agarwal et al., 2017)；当阅读的文本材料连续性和阐述性较高时，进行提取练习是有益的，而当连续性和阐述性较低时，生成性学习更有效，且如果同时还进行提取练习，生成性学习的增益将消失(Roelle & Nückles, 2019)。也就是说，提取练习是不是能在所有情况下，对所有人都能够产生显著增益还存在争议，在实验中用来与提取练习相比较的策略大多是重复学习，未来的研究可以进一步探讨提取练习的适用情况与人群，以及与其他学习策略的比较。

4.2. 临床应用及展望

提取练习除了能够应用到学生的学习上，还能够改善临床人群的记忆力。Wong、Stolwyk 和 Evans (2020)发现提取练习能够显著提高中风患者对于名字的记忆；发育性语言障碍(DLD)的儿童通常在单词学习方面有困难，而提取练习能够改善此类群体的单词学习(Leonard, Deevy, Karpicke, Christ, & Kueser, 2020)。因此提取练习与临床技术应用也有着不小的关联，未来的研究可以进一步探讨提取练习可以改善哪些临床症状，以及如何利用提取练习的系统或工具来帮助改善。

参考文献

张锦坤，张俐娟(2020). 编码与提取时长对提取练习效应的影响. *心理科学, 43(4),* 785-792.

周爱保，马小凤，李晶，崔丹(2013). 提取练习在记忆保持与迁移中的优势效应：基于认知负荷理论的解释. *心理学报, 45(8),* 849-859.

Agarwal, P. K., Finley, J. R., Rose, N. S., & Roediger, H. L. (2017). Benefits from Retrieval Practice Are Greater for Students with Lower Working Memory Capacity. *Memory, 25,* 764-771. https://doi.org/10.1080/09658211.2016.1220579

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering Can Cause Forgetting: Retrieval Dynamics in Long-Term Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1063-1087. https://doi.org/10.1037/0278-7393.20.5.1063

Ariel, R., & Karpicke, J. D. (2018). Improving Self-Regulated Learning with a Retrieval Practice Intervention. *Journal of Experimental Psychology: Applied, 24,* 43-56. https://doi.org/10.1037/xap0000133

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The Mismeasure of Memory: When Retrieval Fluency Is Misleading as a Metamnemonic Index. *Journal of Experimental Psychology: General,, 127,* 55-68. https://doi.org/10.1037/0096-3445.127.1.55

Bertilsson, F., Stenlund, T., Wiklund-Hörnqvist, C., & Jonsson, B. (2020). Retrieval Practice: Beneficial for All Students or Moderated by Individual Differences? *Psychology Learning and Teaching, 20,* 21-39. https://doi.org/10.1177/1475725720973494

Bjork, E. R. A. (1975). The Interaction of Encoding and Rehearsal Processes in the Recall of Repeated and Non-Repeated Items. *Journal of Verbal Learning and Verbal Behavior, 14,* 30-42.

Bjork, R. A. (1988). Retrieval Practice and the Maintenance of Knowledge. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical Aspects of Memory: Current Research and Issues* (Vol. 1, pp. 396-401). New York: Wiley.

Bjork, R. A. (1999). Assessing Our Own Competence: Heuristics and Illusions. In D. Gopher, & A. Koriat (Eds.), *Attention and Performance XVII: Cognitive Regulation of Performance: Interaction of Theory and Application* (pp. 435-459). Cambridge, MA: MIT Press.

Bjork, R. A., & Bjork, E. L. (1992). A New Theory of Disuse and an Old Theory of Stimulus Fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From Learning Processes to Cognitive Processes: Essays in Honor of William K Estes: Vol. 2* (pp. 35-67). Hillsdale, NJ: Erlbaum.

Brewer, G. A., & Unsworth, N. (2012). Individual Differences in the Effects of Retrieval from Long-Term Memory. *Journal of Memory and Language, 66,* 407-415. https://doi.org/10.1016/j.jml.2011.12.009

Butler, A. C., & Roediger, H. L. III (2006). *Feedback Neutralizes the Detrimental Effects of Multiple Choice Testing*. Unpublished Manuscript, St. Louis, MO: Washington University in St. Louis.

Butler, A. C., & Roediger, H. L. III (2007). Testing Improves Long-Term Retention in a Simulated Classroom Setting. *European Journal of Cognitive Psychology, 19,* 514-527. https://doi.org/10.1080/09541440701326097

Carpenter, S. K. (2009). Cue Strength as a Moderator of the Testing Effect: The Benefits of Elaborative Retrieval. *Journal of Experimental Psychology: Learning Memory and Cognition, 35,* 1563-1569. https://doi.org/10.1037/a0017021

Carpenter, S. K., & DeLosh, E. L. (2005). Application of the Testing and Spacing Effects to Name Learning. *Applied Cognitive Psychology, 19,* 619-636. https://doi.org/10.1002/acp.1101

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using Tests to Enhance 8th-Grade Students' Retention of U.S. History Facts. *Applied Cognitive Psychology, 23,* 760-771. https://doi.org/10.1002/acp.1507

Carrier, M., & Pashler, H. (1992). The Influence of Retrieval on Retention. *Memory & Cognition, 20,* 633-642. https://doi.org/10.3758/BF03202713

Chan, J. C. K. (2010). Long-Term Effects of Testing on the Recall of Non-Tested Materials. *Memory, 18,* 49-57. https://doi.org/10.1080/09658210903405737

Collins, A. M., & Quillian, M. R. (1972). How to Make a Language User. In E. Tulving, & W. Donaldson (Eds.), *Organization of Memory* (pp. 309-351). New York: Academic Press.

Coyne, J. H., Borg, J. M., Deluca, J., Glass, L., & Sumowski, J. F. (2015). Retrieval Practice as an Effective Memory Strategy in Children and Adolescents with Traumatic Brain Injury. *Archives of Physical Medicine & Rehabilitation, 96,* 742-745. https://doi.org/10.1016/j.apmr.2014.09.022

Dempster, F. N. (1996). Distributing and Managing the Conditions of Encoding and Practice. In E. L. Bjork, & R. A. Bjork (Eds.), *Human Memory* (pp. 197-236). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-012102570-0/50011-2

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning with Effective Learning Techniques: Promising Directions from Cognitive and Educational Psychology. *Psychological Science in Public Interest, 14,* 4-58. https://doi.org/10.1177/1529100612453266

Fritz, C. O., Morris, P. E., Nolan, D., & Singleton, J. (2007). Expanding Retrieval Practice: An Effective Aid to Preschool Children's Learning. *Quarterly Journal of Experimental Psychology, 60,* 991-1004. https://doi.org/10.1080/17470210600823595

Gardiner, F. M., Craik, F. I. M., & Bleasdale, F. A. (1973). Retrieval Difficulty and Subsequent Recall. *Memory & Cognition, 1,* 213-216. https://doi.org/10.3758/BF03198098

Izawa, C. (1971). The Test Trial Potentiating Model. *Journal of Mathematical Psychology, 8,* 200-224.

https://doi.org/10.1016/0022-2496(71)90012-5

Jacoby, L. L., Bjork, R. A., & Kelley, C. M. (1994). Illusions of Comprehension, Competence, and Remembering. In D. Druckman, & R. A. Bjork (Eds.), *Learning, Remembering, Believing: Enhancing Human Performance* (pp. 57-80). Washington DC: National Academy Press.

Jacoby, L. L., Wahlheim, C. N., & Coane, J. H. (2010). Test-Enhanced Learning of Natural Concepts: Effects on Recognition Memory, Classification, and Metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1441-1451. https://doi.org/10.1037/a0020636

Karpicke, J. D. (2016). A Powerful Way to Improve Learning and Memory. *Psychological Science Agenda.*

Karpicke, J. D., Blunt, J. R., & Smith, M. A. (2016). Retrieval-Based Learning: Positive Effects of Retrieval Practice in Elementary School Children. *Frontiers in Psychology, 7,* Article No. 350. https://doi.org/10.3389/fpsyg.2016.00350

Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive Strategies in Student Learning: Do Students Practise Retrieval When They Study on Their Own? *Memory, 17,* 471-479. https://doi.org/10.1080/09658210802647009

Karpicke, J., & Roediger, H. (2008). The Critical Importance of Retrieval for Learning. *Science, 319,* 966-968. https://doi.org/10.1126/science.1152408

Knight, J. B., Hunter Ball, B., Brewer, G. A., DeWitt, M. R., & Marsh, R. L. (2012). Testing Unsuccessfully: A Specification of the Underlying Mechanisms Supporting Its Influence on Retention. *Journal of Memory and Language, 66,* 731-746. https://doi.org/10.1016/j.jml.2011.12.008

Koriat, A., & Bjork, R. A. (2005). Illusions of Competence in Monitoring One's Knowledge during Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 187-194. https://doi.org/10.1037/0278-7393.31.2.187

Leonard, L. B., Deevy, P., Karpicke, J. D., Christ, S. L., & Kueser, J. B. (2020). After Initial Retrieval Practice, More Retrieval Produces Better Retention than More Study in the Word Learning of Children with Developmental Language Disorder. *Journal of Speech Language and Hearing Research, 63,* 1-14. https://doi.org/10.1044/2020_JSLHR-20-00105

Marsh, E. J. et al. (2009). Memorial Consequences of Answering SAT II Questions. *Journal of Experimental Psychology: Applied, 15,* 1-11. https://doi.org/10.1037/a0014721

Marsh, E. M., Fazio, L., & Roediger III, H. L. (2006). *The Negative Suggestion Effect in Multiple Choice Tests*. Unpublished Manuscript, Durham, NC: Duke University.

Mcdaniel, M. A., & Masson, M. E. (1985). Altering Memory Representations through Retrieval. *Journal of Experimental Psychology Learning Memory & Cognition, 11,* 371-385. https://doi.org/10.1037/0278-7393.11.2.371

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the Testing Effect in the Classroom. *European Journal of Cognitive Psychology, 19,* 494-513. https://doi.org/10.1080/09541440701326154

McDaniel, M. A., Roediger, H. L. I., & McDermott, K. B. (2007). Generalizing Test-Enhanced Learning from the Laboratory to the Classroom. *Psychonomic Bulletin & Review, 14,* 200-206. https://doi.org/10.3758/BF03194052

Meyer, A. N. D., & Logan, J. M. (2013). Taking the Testing Effect beyond the College Freshman: Benefits for Lifelong Learning. *Psychology & Aging, 28,* 142-147. https://doi.org/10.1037/a0030890

Miyatsu, T., & McDaniel, M. A. (2019). Adding the Keyword Mnemonic to Retrieval Practice: A Potent Combination for Foreign Language Vocabulary Learning? *Memory and Cognition, 47,* 1328-1343. https://doi.org/10.3758/s13421-019-00936-2

Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing Learning and Retarding Forgetting: Choices and Consequences. *Psychonomic Bulletin & Review, 14,* 187-193. https://doi.org/10.3758/BF03194050

Pyc, M. A., & Rawson, K. A. (2009). Testing the Retrieval Effort Hypothesis: Does Greater Difficulty Correctly Recalling Information Lead to Higher Levels of Memory? *Journal of Memory and Language, 60,* 437-447. https://doi.org/10.1016/j.jml.2009.01.004

Pyc, M. A., & Rawson, K. A. (2010). Why Testing Improves Memory: Mediator Effectiveness Hypothesis. *Science, 330,* 335. https://doi.org/10.1126/science.1191465

Roediger, H. L. (1978). Recall as a Self-Limiting Process. *Memory & Cognition, 6,* 54-63. https://doi.org/10.3758/BF03197428

Roediger, H. L., & Butler, A. C. (2011). The Critical Role of Retrieval Practice in Long-Term Retention. *Trends in Cognitive Sciences, 15,* 20-27. https://doi.org/10.1016/j.tics.2010.09.003

Roediger, H. L., & Karpicke, J. D. (2006). Test-Enhanced Learning Taking Memory Tests Improves Long-Term Retention. *Psychological Science, 17,* 249-255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roediger, H. L., & Karpicke, J. D. (2006). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science, 1,* 181-210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

Roelle, J., & Nückles, M. (2019). Generative Learning versus Retrieval Practice in Learning from Text: The Cohesion and

Elaboration of the Text Matters. *Journal of Educational Psychology, 111,* 1341-1361. https://doi.org/10.1037/edu0000345

Rowland, C. A. (2014). The Effect of Testing versus Restudy on Retention: A Metaanalytic Review of the Testing Effect. *Psychological Bulletin, 140,* 1432-1463. https://doi.org/10.1037/a0037559

Smith, M. A., Roediger, H. L., & Karpicke, J. D. (2013). Covert Retrieval Practice Benefits Retention as Much as Overt Retrieval Practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1712-1725. https://doi.org/10.1037/a0033569

Szpunar, K. K., Jing, H. G., & Schacter, D. L. (2014). Overcoming Overconfidence in Learning from Video-Recorded Lectures: Implications of Interpolated Testing for Online Education. *Journal of Applied Research in Memory & Cognition, 3,* 161-164. https://doi.org/10.1016/j.jarmac.2014.02.001

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. III (2008). Testing during Study Insulates against the Buildup of Proactive Interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1392-1399. https://doi.org/10.1037/a0013082

Tulving, E., & Arbuckle, T. Y. (1966). Input and Output Interference in Short-Term Associative Memory. *Journal of Experimental Psychology, 72,* 145-150. https://doi.org/10.1037/h0023344

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2011). The Interim Test Effect: Testing Prior Material Can Facilitate the Learning of New Material. *Psychonomic Bulletin and Review, 18,* 1140-1147. https://doi.org/10.3758/s13423-011-0140-7

Wong, D. K., Stolwyk, R. J., & Evans, F. A. (2020). Retrieval Practice Enhances Memory for Names in Survivors of Stroke. *Neuropsychology, 34,* 874-880. https://doi.org/10.1037/neu0000701

Zhang, M., Chen, X., & Liu, X. L. (2019). Confidence in Accuracy Moderates the Benefits of Retrieval Practice. *Memory, 27,* 548-554. https://doi.org/10.1080/09658211.2018.1529796