

记忆提取促进学习: 实验证据与机制解释*

刘兆敏¹ 罗良² 张玮²

(1. 中国政法大学社会学院, 北京 102249;

2. 北京师范大学认知神经科学与学习国家重点实验室, 北京 100875)

摘要: 人类学习的主流观点认为学习主要发生在编码阶段, 记忆提取本身并不产生和促进学习。但最近多项研究对此提出了质疑, 他们发现, 在学习过程中, 记忆提取更能促进学习内容的长时保持。本文首先对记忆提取促进学习的实验证据做简要梳理; 之后重点介绍关于记忆提取促进学习的多种解释机制, 主要包括: 侧重于从理论机制层面解释的记忆失用新理论, 强调内在精细机制的精细提取假设, 以及解释保持间隔(长、短)与学习方式(重复学习、学习-测试)交互作用的基于分布的二分模型; 文章最后分析了当前研究记忆提取作用实验存在的一些问题, 提出了今后的研究方向, 强调应从多角度揭示记忆提取练习促进学习的机制。

关键词: 记忆提取练习效应; 记忆失用新理论; 精细提取假设; 基于分布的二分模型

长期以来, 如何提高学生的学习效率这一棘手难题备受国内外心理学研究者和教育实践者的关注。大多数的心理学者和教师将目光聚焦于能促进学生对信息进行有效“编码”的方法上, 即寻找能把知识有效输送到长时记忆中的策略, 例如分散学习 (spacing learning) 与集中学习 (mass learning) 的研究, 过度学习 (overlearning) 的研究等。但是, 同样作为记忆重要加工阶段的“提取”在促进学习方面的积极作用却一直很少受到研究者的关注 (Karpicke & Roediger, 2008)。出现这种局面的主要原因在于, 关于人类学习和记忆本质的主流假设认为学习主要发生在编码阶段, 记忆提取就像测量一个物理客体的大小、形状和重量一样, 只是用来测量先前学习效果的手段, 其本身并不改变记忆, 更不能产生和促进学习 (Karpicke & Blunt, 2011)。但是, 最近大量行为研究的结果对这一观点提出了质疑, 并指出, 与简单重复学习和过度学习相比, 学习过程中进行一次或几次记忆提取更能显著促进学习内容的长时保持 (Roediger & Karpick, 2006a; Karpicke & Roediger, 2008; Karpicke & Blunt, 2011; Carpenter, 2009; Pyc & Rawson, 2009; Rohrer, Taylor, & Sholar, 2010)。研究者将学习过程中一次或多次记忆提取导致最后测试成绩提高的现象称为提取练习效

应 (retrieval practice effect), 也称测试效应 (testing effect)。本文将主要对记忆提取促进学习的实验证据、机制解释及相关研究可能存在的问题进行梳理。

1 记忆提取促进学习的实验证据

早在两千多年前, 希腊哲学家亚里士多德就提出: 重复练习回忆一件事情可以增加对它的记忆。但是直到 1909 年, Abbott 使用无意义音节和单词作为实验任务, 才第一次发现记忆提取促进学习的实验证据。然而在之后长达 80 多年时间里, 只有少数研究关注记忆提取促进学习现象 (Gates, 1917; Spitzer, 1939), 大部分学习和记忆实验, 只把记忆提取作为测量学习效果或者不同编码条件效果的手段, 并对不同条件下的记忆提取方式进行严格控制, 以防止对实验结果产生干扰 (Roediger & Butler, 2011)。进入 21 世纪, 记忆提取促进学习的研究逐渐丰富, 相关成果陆续发表在 *Science*, *Psychological Science* 等重要杂志上 (Karpicke & Roediger, 2008; Karpicke & Blunt, 2011; Pyc & Rawson, 2010; Roediger & Karpick, 2006)。这些研究使用严密的实验设计, 为揭示记忆提取促进学习现象提供了更为客观的证据。

2006 年, Roediger 和 Karpicke 在一项研究中,

* 基金项目: 国家自然科学基金资助项目 (批准号: 31000506); 教育部人文社会科学研究青年基金项目 (批准号: 10YJCXLX032); 中国政法大学校级人文社会科学研究资助项目。

通讯作者: 罗良, E-mail: luoliang@bnu.edu.cn

让所有被试用两种学习方式在相等的时间内学习科学小品文,一种方式是重复学习,即学习(S)-再次学习(S),简称SS条件;另一种是学习(S)-测试(T)方式,简称ST条件,测试时要求被试尽可能回忆文章的内容并写下来(对回忆后的内容不给予任何反馈)。这一阶段结束之后,被试随机分为三组,每组被试在不同的时间间隔后进行记忆测试:5分钟、2天和7天。研究者发现(请见图1中的A图),5分钟之后测试在SS条件下的成绩显著高于ST条件,但在2天和7天之后测试中,ST条件却都显著高于SS条件。在这一研究基础上,Roediger及其同

事又设计了第二个实验,把学习方式分为学习-学习-学习-学习(即重复学习4次,简称SSSS),学习-学习-学习-测试(即重复学习3次加1次记忆测试,简称SSST)和学习-测试-测试-测试(即学习1次加3次记忆测试,简称STTT)三种条件,分别在5分钟和7天之后进行测试(请见图1中的B图),发现在5分钟之后的测试中,测试结果为SSSS > SSST > STTT,但7天之后的测试结果完全相反,即STTT > SSST > SSSS。即,与简单的重复学习相比,提取更有利于记忆内容的长时保持。

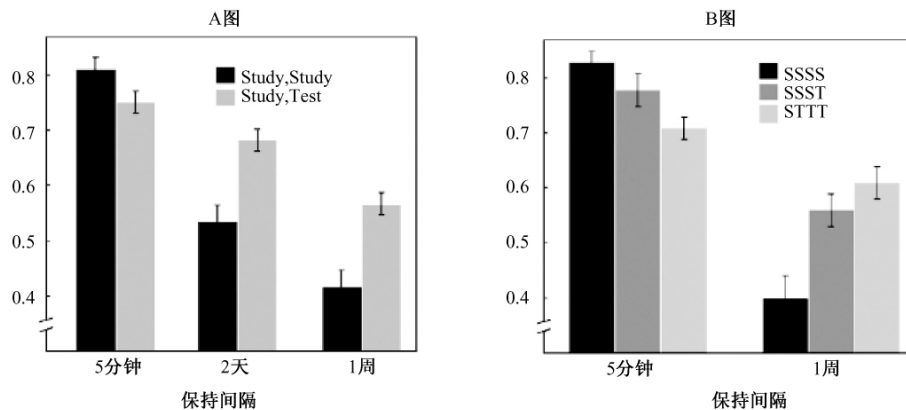


图1 Roediger实验的数据结果图(2006a)

在Roediger和Karpicke(2006)的研究中,只是把记忆提取与简单的重复学习进行了对比,那么记忆提取与强调精细化编码的精加工学习相比,是否还占有优势?Karpicke对此也进行了探讨,他在Science(2011)杂志上发表的一项研究发现,即使与要求对知识进行精细加工和编码的概念图(Concept mapping)学习相比,记忆提取在促进学习内容的长时保持方面仍占有优势。多项使用不同记忆材料的实验,如单词表(Carpenter & DeLosh, 2006)、外语单词(Karpicke & Roediger, 2008; Pyc & Rawson, 2009; Toppino & Cohen, 2009)、一般事实性知识(Carpenter, Pashler, Wixted, & Vul, 2008)、文本段落(Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008)、姓名面孔匹配(Carpenter & DeLosh, 2005)、视觉空间信息(Carpenter & Pashler, 2007; Rohrer, Taylor, & Sholar, 2010)等,也都证实了提取练习效应的存在。这些研究发现,提取练习效应与学习过程中的测试是否提供结果反馈没有必然关系,即在学习阶段进行没有结果反馈的提取练习,仍然能够促进学习信息的长时保持。

2 记忆提取促进学习的机制解释

记忆提取练习为什么能够促进学习?其背后的机制是什么?早期研究者利用记忆编码促进学习的思路来解释这一现象,认为提取可以让学习者重新接触到学习材料,引起被成功提取材料的过度学习,因此导致了测试成绩的提高(Thompson, Wenger, & Bartling, 1978)。但是前面提到的多个研究表明,对提取练习与重复学习的时间和次数进行了严格匹配,仍然发现了明显的提取练习效应,否定了这种解释。此外,还有研究者试图利用迁移适当加工理论(transfer-appropriate processing)来解释记忆提取练习效应,他们认为出现提取练习效应的原因在于,与重复学习相比,学习过程中的初次测试创造了与最后测试相似的条件,从而引起了提取练习效应。他们以此推测初次测试与最后测试之间的相似性越高,最后测试的成绩就会越好。但是多个实验结果并没有支持这种解释(Carpenter & DeLosh, 2006; Kang, McDermott, & Roediger, 2007)。例如,在Carpenter和DeLosh(2006)的实验中,初次测试和最后测试各有三种类型:再认、线索回忆和自由回

忆,有的实验条件下初次测试与最后测试类型相似(再认-再认),但有的条件下不同(再认-自由回忆)。实验结果并不像迁移适当加工理论预测的那样,初次测试与最后测试相似性越高,最后测试成绩越好。结果却发现,不管两者条件是否相似,初次测试为自由回忆方式时,最后测试时对项目的记忆成绩最好。因此,目前研究者逐渐将提取练习效应的机制解释聚焦于提取本身是如何影响记忆的,即提取与编码相比,对学习产生什么样的特殊影响,并提出了不同的理论观点。

2.1 记忆失用新理论的解释

要真正理解提取练习效应产生的本质,必须对人类记忆本身的特性有更深入的认识,Bjork夫妇(1992)提出的记忆失用新理论(A new theory of disuse)在这方面进行了尝试。该理论认为,人类记忆系统具有与人造记忆系统(例如录音带或者计算机)不同的特性:(1)新信息存储到记忆中的过程并不是对这些信息的精确复制,而是根据已知信息对这些新信息进行的解释。即根据它们的意义,将新信息“装配”到已有记忆中,这种加工似乎没有容量限制;但当需要使用这些信息时,提取加工则具有高度的线索依赖性,并且不稳定,容易出现错误。例如,在一个情景下,非常容易提取出来的信息(姓名、电话号码或者街道名称),在另外一个情景下可能提取不出来。因此能否成功提取信息取决于能够获取的线索,这些线索包括各种环境的、人际的、情感的或者生理的信息。(2)从记忆中提取信息会修改已有记忆系统,被成功提取的信息其可提取性将增强,而没被成功提取的信息其可提取性将会降低,即提取本身是一个潜在的学习事件,不仅对提取的信息有影响,对没被成功提取的信息也会产生影响。

Bjork夫妇基于对人类记忆特性的分析,提出记忆中的每个项目均具有两个“强度”特征:存储强度(storage strength)和提取强度(retrieval strength)。存储强度是指一个项目学的怎么样(即记忆痕迹的相对持久性),提取强度则是指项目在当前情景下的可通达性。记忆中一个项目被回忆出的可能性只取决于与它对应的线索的提取强度,而与存储强度没有直接关系。即,存储强度只是一个对成绩没有直接影响的潜在变量。而且提取强度与存储强度没有必然的对应关系,当一个存储强度较弱的项目近期被频繁提取时,其提取强度将增强,可能会超过很多其他的具有高存储强度的项目。例如,你会非常清楚地记住当前一个仅住了3天的旅馆房间号,但

可能想不起20年前持续用了5年的电话号码。Bjork夫妇的理论还假设,学习一个项目和从记忆中提取一个项目都能增加该项目的提取强度和存储强度,但是提取导致的增加量更大。

Bjork夫妇(1992)提出的记忆失用新理论强调了提取在修改已有记忆系统中的重要作用,这为提取练习效应提供了理论基础。该理论提出的与学习相比,从记忆中成功提取一个项目能够导致更大的提取强度和储存强度增加量,可以很好地解释与简单重复学习相比,为什么提取练习能够导致最后测试成绩更好的现象。

2.2 精细提取假设的解释

Bjork夫妇(1992)从理论层面对记忆提取促进学习的机制进行了探讨,但是并没有揭示提取促进学习的精细机制,而且提取强度的概念也非常抽象。为解释提取加工本身能够直接促进学习的精细化机制,Carpenter(2009)在Collins记忆的激活扩散模型基础上提出了精细提取假设(Elaborative retrieval hypothesis)。她认为,与SS条件下再次学习时信息都呈现出来相比,ST条件下的提取测试只呈现提取线索(或不呈现任何信息),被试需要在长时记忆中努力搜索没呈现出来的、之前学习过的内容,在这种搜索过程中,可能会激活多个与学习内容相关的概念,并可能形成一个新的语义概念网络,或使原有语义概念网络间的联结通路增强。以“面包-篮球”(篮球为目标词)为例,在提取测试时,要求被试看到面包便回忆出篮球。被试在搜索目标词的过程中,可能会激活多个概念,并有可能建立起“面包-鸡蛋-圆形-篮球”的语义概念网络,这种建立过程本身就是一种精细化的加工模式,而根据这种模式建立起来的网络将对信息的长时保持有利。提取练习之所以更能促进学习,就在于提取比重复学习更有可能形成精细化的语义概念网络。因此,她认为是精细化的语义概念网络导致了提取练习效应。

Carpenter(2009)通过实验对精细提取假设进行了证明。她在实验中让被试学习不同语义相关程度的词对(一个作为线索词,另一个为目标词),学习方式分为SS(重复学习)和ST(学习-测验)两种。在两种学习方式的学习阶段,要求被试在5点量表上判断两个词之间的相关程度,并进行记忆。在ST条件下,被试在测试阶段要根据给出的线索词回忆目标词(线索提取测试),但不反馈测试结果。SS和ST结束的5分钟后,要求被试回忆出所有的目标词(自由回忆测验)。Carpenter预计,如果SS不产生

精细化语义概念网络的激活,那么,高语义和低语义相关词对在自由回忆测验上的成绩应该没有显著差异;而对于 ST 来说,由于低语义相关词对在线索提取阶段可能会激活更多精细化的语义概念,因此在自由回忆测验上的成绩可能显著好于高语义相关词对。她的实验结果也的确证实了这一假设。Chan, McDermott 和 Roediger(2006) 也发现,在学习过程中被试对部分学习内容提取练习能促进那些没接受提取练习、但与提取练习内容相关联信息的学习,间接支持了精细提取假设。

Psy 和 Rawson(2010) 对精细提取假设进行了发展,提出了中介效应假设(Mediator effectiveness hypothesis)。中介效应假设认为,提取练习促进学习的原因在于,记忆测试支持在编码阶段使用了更有效的中介物,这个中介物可以是一个把线索和记忆目标联系在一起的单词、短语或者概念。中介物的作用大小取决于先后进行的两个加工:中介提取(mediator retrieval)和中介解码(mediator decoding)。中介提取是指当提取线索信息时,中介物能够被回忆出来;而中介解码则是在中介物能成功回忆的基础上,把与中介物联系的记忆内容提取出来,这一假设认为记忆测试条件下中介物更有可能被提取,其解码能力更强。Psy 和 Rawson 让被试学习英语与斯瓦希里语匹配的词对,其中英语单词是线索词,斯瓦希里语是目标词。实验有 SS 和 ST 两种条件,要求被试在每个 trial 中都要生成并口语报告出一个中介物(关键词),要求这个关键词在外形或读音上与斯瓦希里语单词相似,但意义要与英语单词相关。最后的测试在一周后进行,有三种条件:单独线索词(只出现线索词,让被试回忆目标词)、线索词加中介词(出现线索词和中介词,让被试回忆目标词)、线索词加回忆中介词(出现线索词,先让被试回忆中介词,然后再回忆目标词)。实验结果支持了中介效应假设。即,研究者发现,在线索词加回忆中介词条件下,ST 条件回忆出中介词的比例显著高于 SS 条件,说明记忆测试促进了中介提取;而且,成功回忆出中介词后进一步回忆出目标词的比例方面,ST 条件显著优于 SS 条件,证明了 ST 条件下的中介解码能力更强。

2.3 基于分布的二分模型的解释

以上介绍的记忆失用新理论和精细提取假设都不能对以往多个研究发现的保持间隔(长、短)和学习方式(重复学习、学习-测试)间的交互作用进行解释,即为什么在短间隔条件下(学习结束 5 分钟

后进行最后测试),SS 的最后测试成绩优于 ST,但在长间隔条件下(学习结束 2 天或者 7 天后进行最后测试),ST 的最后测试成绩优于 SS。研究者提出了基于分布的二分模型(Distribution-based bifurcation model)对此进行了回答(Halamish & Bjork, 2011; Kornell, Bjork, & Garcia, 2011)。基于分布的二分模型以 Bjork(1992)提出的项目记忆强度均具有存储强度和提取强度为理论基础,并从项目提取强度分布变化的角度对提取练习促进学习的机制进行解释的。其基本假设是:不同项目的记忆强度分布是一个连续体,当某个项目当前的记忆强度超过回忆阈值时,这个项目就能被回忆出来。因此,回忆测试并不是对记忆强度的直接测量,而是考察了项目记忆强度是否超过了回忆所要求的阈限。从这一基本假设出发,该模型认为(见图 2 所示),初次学习之前,对于所有被试和所有要学习的项目,其提取强度都呈正态分布。初次学习之后,所有项目的提取强度都得到了增加,而且仍呈正态分布。但是,当一些记忆项目经过记忆提取练习阶段后,他们的记忆强度分布变的将不再是正态分布。假设存在两个记忆项目列表 A 和 B,它们都经过了初次学习,在接下来的实验中,列表 A 中的项目进行没有结果反馈的提取练习,而列表 B 中的项目则进行与初次学习相似的再次学习。对于列表 A 来说(见图 2 中右侧图),经过提取练习后,部分项目被成功提取,这些项目的记忆强度有比较大的增加,而没有被成功提取的项目其记忆强度没有增加,并且会越来越弱。因此,通过提取练习,列表 A 中的项目被分成了两部分,不再是一个连续的正态分布,被成功提取的项目记忆强度明显增强;而提取失败项目的记忆强度不仅没有增加反而减弱了。而列表 B 中的项目,经过再次学习之后,所有项目都得到了不同程度的增强,仍然呈连续的正态分布。

基于分布的二分模型认为,最后测试的难度会调节记忆提取练习效应的大小,这也是出现保持间隔(长、短)和学习方式(重复学习、学习-测试)之间交互作用的原因。如图 2 所示,带箭头的蓝色竖线表示测试难度。在 ST 条件下,如果项目测试成功提取率为 57%,那么只有 57% 的项目的记忆强度被显著加强了,而在 SS 条件下,所有项目记忆强度都被加强了,但是记忆强度增加幅度上比 ST 条件下被成功提取的项目要小。短保持间隔后进行最后测试(假设该测试与 ST 条件中初次测试的绝对难度相同),SS 条件下可能有 70% 的项目能被成功提取,

但是 ST 条件则最多只能有 57% 的项目被成功提取, 这时的 SS 条件的最后测试成绩要优于 ST 条件。但是, 如果经过长时间保持间隔后进行最后测试 (仍假设该测试与 ST 条件下初次测试的绝对难度相同), 无论对于 SS 条件还是 ST 条件中的项目, 其项目记忆强度都降低了, 但是由于 SS 条件下, 其项目记忆强度比 ST 条件在初次测试中被成功提取的项目要低, 所以, SS 条件下, 将有更多项目的记忆强度变得低于最后测试要求的记忆强度。因此, 长保

持间隔后, ST 的最后测试成绩优于 SS 条件。这个模型也可以很好地解释最后测试任务为再认的实验中, 为什么没有发现记忆提取的练习效应: 再认任务的难度较低, 即使是长保持时间间隔, SS 条件下仍有很多项目记忆强度高过再认任务要求的记忆强度, 因此, SS 条件的最后测试成绩与 ST 条件没有显著差异, 甚至高于 ST 条件。Kornell, Bjork 和 Garcia (2011), Halamish 和 Bjork (2011) 在实验中通过创设不同测试难度的任务, 其结果也支持了这一模型。

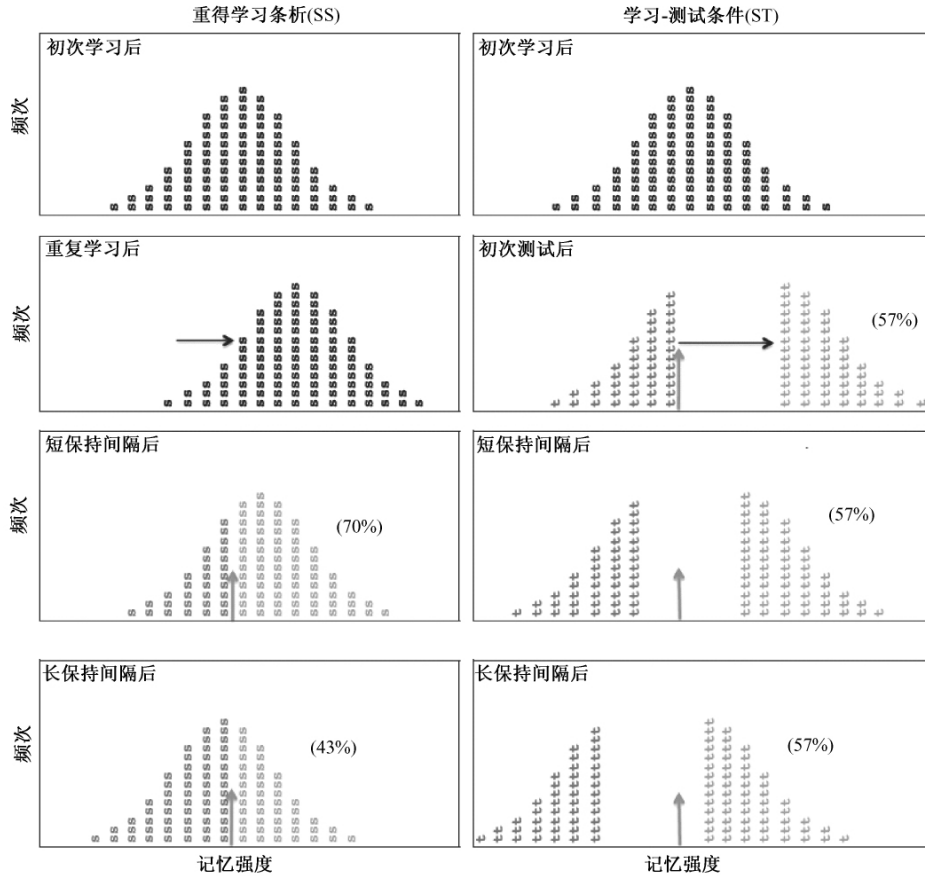


图2 基于提取强度分布的二分模型 (Kornell, Bjork, & Garcia, 2011)

3 需要继续探讨的问题

综上所述, 关于提取练习能够促进学习的现象已在大量研究中得到证实, 研究者提出的假设和理论从不同侧面对其机制进行了解释。失用新理论更多从理论层面提供解释, 其重要贡献是区分了存储强度和提取强度; 精细提取假设的重要突破是从精细化的加工层面揭示了记忆提取为什么能够促进学习, 其实质是解释了记忆提取练习提高项目提取强度的内部机制; 而基于分布的二分模型的聚焦点并不在于揭示记忆提取促进学习的内在机制, 而是为了解释保持间隔(长、短)和学习方式(SS、ST)出现

交互作用的原因。但是关于提取练习促进学习的研究总体上还处在初级阶段, 目前还存在以下问题需要在将来的研究中进一步解决:

第一, 目前关于提取练习能够促进学习的研究结论基本上全是由行为数据得来的, 需要脑机制研究方面的证据提供支持。例如, Carpenter (2009) 发现在 ST 中低语义相关条件下根据线索词提取目标词的反应时显著长于高语义相关条件下的反应时, 并进而推测低语义相关条件可能激活了更精细化的语义网络, 而这完全可以在大脑动态活动模式上进行探讨。Roediger (2011) 对近十几年提取练习促进学习的实验进行综述之后也提出, 关于提取练

习促进学习的理论本质亟需脑机制层面的研究。

第二,在以往关于提取练习促进学习机制的研究中,实验材料主要以配对联想词对为主,其他类型的刺激为学习任务的实验还比较少。而现实生活中的学习有各种材料,如图片、散文、各种技巧等,虽然部分实验已经揭示记忆提取练习可以促进这些材料的学习,但是其机制是否与以配对联想词对为实验任务所揭示的机制一致,目前并没有答案,认识这些机制对于完善提取练习促进学习的机制具有重要意义。

第三,以往的研究主要以大学生的学习为前提,从儿童发展和教育实践的角度进行的研究还不多。例如,Rohrer, Taylor, Sholar(2010)以4年级小学生作为被试,发现了提取练习效应的存在;张锦坤(2008)对小学五年级、初二和高二学生的提取练习效应进行了研究,但只是在每个年级进行了对比,并没有比较这种提取练习效应的发展特点。那么,更低年级的儿童是不是存在提取练习效应?什么形式的提取练习更适合儿童青少年?如何根据提取练习效应来提高学生的学习效果?如何合理利用考试促进儿童青少年的学习?怎样利用记忆提取指导教学实践?这些都是需要迫切解决的科学和实践问题。

参考文献:

- Abbott, E. E. (1909). On the analysis of the factors of recall in the learning process. *Psychological Monographs*, 11(1), 159 - 177.
- Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger, H. L., III, & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology*, 22, 861 - 876.
- Bjork, R. A., & Bjork, E. A. (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.
- 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(6), 1563 - 1569.
- Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition*, 34, 268 - 276.
- Carpenter, S. K., & Pashler, H. (2007). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review*, 14, 474 - 478.
- Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, 19, 619 - 636.
- Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition*, 36, 438 - 448.
- Chan, J. C. K., McDermott, K. B., & Roediger, H. L. III. (2006). Retrieval induced facilitation: Initially non-tested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General*, 135, 553 - 571.
- Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of psychology*, 6, 1 - 104.
- Halamish, V., & Bjork, R. A. (2011). When does testing enhance retention? A distribution-based interpretation of retrieval as a memory modifier. *Journal of Experimental Psychology: Learning, Memory, and Cognition* (in press).
- Kang, S. H. K., McDermott, K. B., & Roediger, H. L., III. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, 19, 528 - 558.
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. *Science*, 331, 772 - 775.
- Karpicke, J. D., & Roediger, III. H. L. (2008). The critical importance of retrieval for learning. *Science*, 319, 966 - 968.
- Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: a distribution-based bifurcation model. *Journal of Memory and Language* (in press).
- 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.
- Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: mediator effectiveness hypothesis. *Science*, 330, 335.
- Roediger, III. H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Science*, 15(1), 20 - 27.
- Roediger, III. H. L., & Karpicke, J. D. (2006a). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249 - 255.
- Roediger, III. H. L., & Karpicke, J. D. (2006b). The power of testing memory: basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181 - 210.
- Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(1), 233 - 239.
- Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology*, 30, 641 - 656.
- Thompson, C. P., Wenger, S. K., & Bartling C. A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning and Memory*, 4(3), 210 - 221.
- Toppino, T. C., & Cohen, M. S. (2009). The testing effect and the retention interval: Questions and answers. *Experimental Psychology*, 56(4), 252 - 257.

Retrieval Practice Enhance Learning: Experimental Evidence and Underlying Mechanism

LIU Zhao-min¹ LUO Liang² ZHANG Wei²

(1. School of Sociology, China University of Political Science and Law, Beijing 102249, China;

2. State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China)

Abstract: Most thought on human learning is that learning happens primarily when people encode knowledge, and retrieval does not itself produce learning. However, a number of studies have found that retrieval produces more learning than repeated study, which challenges the dominant assumption. In this article, we introduce the experimental evidence that retrieval practice enhance learning. And there are several theories of revealing the beneficial effects of retrieval practice: (1) the new theory of disuse, which focuses on the theoretical explanations for retrieval practice effect; (2) elaborative retrieval hypothesis, which emphasizes fine underlying mechanism of retrieval practice effect; (3) a distribution-based bifurcation model, which explains the test-delay interaction of study on retrieval practice effect. Finally, we will discuss some deficiencies in previous studies on retrieval practice effect, and future investigation should be strengthened by using different materials in multiple aspects.

Key words: retrieval practice effect; the new theory of disuse; elaborative retrieval hypothesis; distribution-based bifurcation model