

利用测试促进学习:记忆心理学的研究进展与教育启示

罗 良,张 玮

(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】 记忆心理学研究发现,测试不仅是评价先前学习效果的手段,它还可以改变长时记忆中的信息。与简单重复学习相比,学习过程中进行一次或几次不给予任何反馈的测试能显著促进学习内容的长时保持。测试对学习的促进获得了多方面实验证据的支持,对此研究者也提出了不同的理论解释,包括侧重于理论机制的记忆失用新理论和强调内在机制的精细提取假设。影响测试促进学习效果的因素则包括了测试次数、测试间隔时间、测试形式及测试后反馈等四个方面。在当前教学实践中,需要重视测试本身对学习的促进作用,如教师要学会合理利用课堂测试,培养学生学会和使用自我测试的学习策略,在设计教材与各类辅导书中的习题以及开发多媒体学习软件时要充分利用测试促进学习研究的发现等。

【关键词】 测试效应;记忆提取;学习;记忆心理学

【中图分类号】 B84 **【文献标识码】** A **【文章编号】** 1002-0209(2012)01-0043-08

测试在学校教育中扮演着重要角色,是教学中考察和评估学生知识掌握情况的一种必不可少的方式。通过测试,可以鉴定学生的能力水平和学业成绩,诊断学业困难以及决定升学、毕业、升级、留级、班级划分等。测试在记忆心理学研究者看来是一个个体从长时记忆中提取信息的过程,即记忆提取。长期以来,困扰他们的问题是:测试中发生的记忆提取难道真像我们平时测量一个物理客体的大小、形状和重量一样,只是用来评价先前学习效果的一种手段吗?测试过程本身可不可以改变已存储在长时记忆中的信息,可不可以促进这些信息的保持?一百多年来,记忆心理学研究者围绕着这些问题进行了丰富的研究,尤其是最近10年的研究发现,测试不仅是评价先前学习效果的手段(用来考察有多少信息保持在了长时记忆中),它还可以改变长时记忆中的信息;而且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测试更能显著促进学习内容的长时保持,即使对测试不给

予任何反馈,仍能显著促进学习效果,研究者将这种现象称为测试效应(testing effect),也叫提取练习效应(retrieval practice effect)(Butler, Karpicke, & Roediger III, 2007; Carpenter, 2009; Karpicke & Blunt, 2011; Karpicke & Roediger III, 2008; Pyc & Rawson, 2010; Roediger III & Karpicke, 2006b)。由于这一效应的教育应用对于提高学生学习效率具有重要价值,引起了心理学研究者和教育实践者的广泛关注,本文首先介绍测试促进学习的实验证据和相关理论解释,接下来详细阐述影响测试促进学习效果的可能因素,最后分析这些研究结果对于当前教育的启示。

一、测试促进学习的实验证据和理论解释

1909年,Abbott使用无意义音节和单词作为实验材料,第一次发现了测试促进学习的实验证据。Gates(1917)和Spitzer(1939)分别使用大样本儿童青少年被试发现测试能够促进学习,然而这

【收稿日期】 2011-11-07

【基金项目】 国家自然科学基金资助项目“记忆提取促进学习的机制研究”(31100817)、“基于激活的模型中注意和可用的长时记忆间的关系”(31000506);教育部人文社会科学研究青年基金项目“工作记忆的分子遗传学研究”(10YJCXLX032)

些实验在当时并没有引起研究者和教育工作者的注意。直到1967年,记忆心理学的奠基者之一、著名加拿大心理学家 Tulving(1967)的研究引发了人们去思考重复测试和重复学习对记忆的不同影响。他让被试在三种不同的学习条件下学习36个词语,分别是:标准学习条件(学习—测试—学习—测试,即STST条件),重复学习条件(学习—学习—学习—测试,即SSST条件),重复测试条件(学习—测试—测试—测试,即STTT条件),并且在测试时不给予结果反馈。学习结束后要求被试自由回忆词语,发现三组的差异并不显著:STST条件和SSST条件下都回忆出约20个词语,STTT条件下回忆出18.5个词语;三组的学习曲线也保持较高的一致性。这表明测试不仅仅能考察学习效果,测试也能促进学习。

Tulving的研究证明了测试对学习的促进作用,但其最终测试是学习完成之后即刻进行的,并没有考察测试是否对长期学习有促进作用。2006年,Roediger和Karpicke以科技散文作为实验材料,研究了测试对信息长时保持的影响。在实验中,研究者设计了两种学习方式:重复学习(学习—学习,简称SS条件)、学习测试(学习—测试,简称ST条件),测试时要求被试回忆文章内容(对回忆后的内容不给予任何反馈)。这一阶段结束之后,两种学习条件下的被试被随机分为三组,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(见图1),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,ST条件却都显著高于SS条件。进而证明了测试比重复第二次学习更有利于学习的长期保持。在Roediger和Karpicke(2006b)的研究中,只是把测试条件与简单重复学习条件进行了对比,那么当测试与强调精细加工的重复学习方式相比较时,还能表现出优势吗?Karpicke在*Science*上发表的一项研究对此进行了探讨(Karpicke & Blunt, 2011),他们发现,即使与要求对知识进行精细加工编码的概念图(Concept mapping)学习方法相比,测试在促进学习内容的长时保持方面仍占有优势。大量实验发现了测试对学习内容的长时保持的促进作用,Rawson和Dunlosky(2011)在一篇文章中对过去100多年中测试促进学习的研究进行了综述,至少有150多篇文章中的300多个实验发

现了测试效应,2000至2010年的10年间就有82篇文章和168个实验之多,其中多篇文章发表在*Science*(2008;2010)等顶级杂志上。

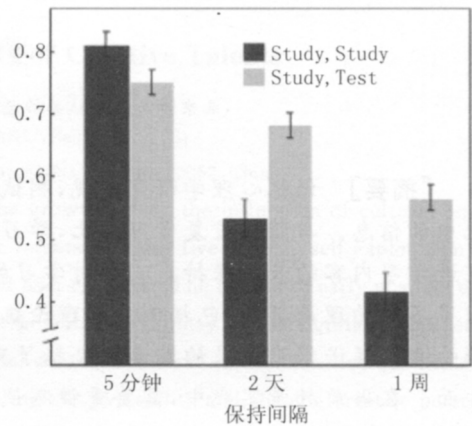


图1 Roediger和Karpicke实验的数据结果图(2006)

测试为什么更能促进学习?其理论机制是什么?Roediger和Karpicke(2006a)将测试对学习的促进作用分为间接效应和直接效应。所谓间接效应是指由于测试引发了其他一些加工过程发生了变化,从而促进了学习,例如通过有反馈的测试,学生可以知道自己知识掌握的状态,从而指引其在未来学习中合理分配时间等。而直接效应则是指测试本身对学习和信息长时保持的促进作用,即与重复学习相比,学习时间和接触学习材料机会上没有增加,没有给予额外结果反馈,仅仅通过测试本身促进了学习和信息长时保持。关于测试促进学习理论机制的探讨,主要集中在直接效应上。

Bjork和Bjork(1992)的记忆失用新理论(new theory of disuse),从记忆特征角度对测试本身促进学习的理论机制进行解释。该理论认为,与人工记忆系统(例如录音带或者计算机)不同,人类记忆系统里面的信息具有动态变化的特性。个体在存储新信息时并不是简单地对信息进行复制,而是会根据长时记忆中的已有信息对新信息进行解释、建构;提取信息也不仅仅是读取的过程,信息被成功提取后它的存储状况会发生变化,变得更容易提取,提取是“记忆的修改器”(memory modifier)。根据人类记忆的这些特性,Bjork提出记忆中的每个项目均有两个“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。提取强度具有高度的线索依赖性,即信息能否成功提取取决于特定情境下可获取的线索,而与存储强度没有直接关系。一个项目的当前提取强度高并不意味着存储强度也高,而存储强度高的项目也会因为缺乏可利用的线索而降低了提取强度。基于记忆失用新理论,Bjork(1994)提出了适度学习难度(desirable difficulties)原理来解释测试为什么比重学更能促进学习。他指出,要想让知识长久地保持下去,获得较高的存储强度,必须要增加学习时的难度,因为困难会迫使人们努力对项目进行更深入的加工,从而记忆痕迹也会更深。增加学习难度的方法之一就是使用测试来代替重学,因为测试需要人们主动提取信息,这比被动重学的难度要大,因此测试导致的增加量更大,促进了对该项目的持久记忆。

Bjork的理论提供了一个具有较广解释力的假设,但没有描述测试促进记忆的精细机制。为此,Carpenter(2009)提出了精细提取假设(Elaborative retrieval hypothesis)来解释测试促进学习的精细加工机制。该假设认为与重学条件相比,测试条件下的提取会更多地激活没有呈现出来的、但与学习内容相关的其他概念,从而形成一个新的语义网络或者使已有语义网络的联结更强更丰富,这个精细的语义网络使得长时间间隔后的提取路径更加多元化。Pyc和Rawson(2010)进一步拓展了精细提取假设,提出了中介效应假设(Mediator effectiveness hypothesis)。中介是指把线索和记忆目标联结在一起的词、短语或者概念等。从线索经过中介提取记忆目标包含两个加工过程:中介提取(mediator retrieval)和中介解码(mediator decoding)。中介提取是指当呈现线索时,将中介物回忆出来;而中介解码则是指成功提取出中介后,把与中介物联系的记忆目标提取出来,Pyc & Rawson(2010)和Carpenter(2011)的实验发现,测试条件下中介物可提取性更强,其解码能力更强,支持了该假设。

二、影响测试促进学习效果的因素

在寻求测试促进学习的实验证据与理论解释的同时,记忆心理学研究者还对多个可能影响测试促进学习效果的因素进行了研究,这些影响因素主要包括:测试次数、测试间隔时间、测试形式以及测

试的结果反馈等。

(一)测试次数

既然测试能够促进学习,那么增加测试的次数,学习内容会不会保持得更久?Wheeler和Roediger(1992)让被试记忆60个实物名称,第一组被试在记忆结束后即离开,第二组被试进行一次自由回忆,第三组被试进行三次自由回忆,即三组接受测试的次数分别是0次、1次、3次。一周后要求所有被试回忆实物名称,第一、二、三组的成绩分别是29%、39%、53%,遗忘率分别是46%、27%、13%。可见随着测试次数的增加,被试记得更牢,遗忘得更慢。其他多位研究者也发现了相似的结果(Allen, Mahler, & Estes, 1969; Wheeler & Roediger, 1992; Roediger III & Karpicke, 2006b)

当然,有的研究发现测试次数也并不是越多越好。Rawson和Dunlosky(2011)就考察了这样一个问题:要想获得持久、有效的学习,学习阶段进行多少次成功提取最合适。他们让被试阅读一篇记忆心理学文章,其中包括一些基本概念如感觉记忆、情景记忆,阅读之后要求被试自由回忆这些概念,每次回忆后均提供反馈。他们控制了成功回忆的次数:被试需要成功地自由回忆出每个概念1次、2次、3次或4次。在2天后所有被试又进行了有反馈的测试,要求所有概念达到成功自由回忆一次,即相当于对每个概念复习了一次。6周后所有被试接受一次线索回忆的最终测试。他们发现在2天后对所有概念的初次自由回忆中,之前成功回忆2次的比只成功回忆1次的要好,成功回忆3次的也比2次的要好,但是成功回忆出4次的和3次的两者之间没有显著差异。可见一开始随着测试次数增多,被试也能记得更多内容。但当次数增加到一定程度时,其表现就不再继续提高了。该实验还发现经过2天后的有反馈测试,6周后所有被试对这几个概念的掌握程度几乎相同了(四组被试之间均无显著差异),表明初学时重复提取的优势在经过一次2天后的有反馈测试后消失了。Rawson等人指出虽然初学时重复提取能增强记忆,但它耗费了更多时间:初学时多测3次和在间隔2天后多测一次的效果几乎相同,因此在学习之后间隔一段时间进行复习是非常高效的。不过这里并不是否认初学时多次成功提取的必要性,而是说每个人应当根据自己的情况来安排学习和复习时的测试:如

果能够进行复习测试的话,初学时可以不必有太多测试;但对于大部分学生来说,让他们记得在2天后安排一次复习测试实在是太难了,那么初学时就必须要多提取几次了。

(二)多次测试的间隔时间

前面提到合理安排多次测试比一次测试更能促进记忆,但多次测试又有多种编排方法,例如可以是集中测试,也可以是分散测试;分散测试又包括相等时间间隔的测试,递增时间间隔的测试等。那么如何安排多次测试才能产生最好的效果呢?根据间隔效应(spacing effect),分散的练习比集中的练习更有利于知识的长期保持,因此分散测试比集中测试在促进长时记忆上的效果要好。这一点在很多考察测试间隔的研究中也得到了验证(Cull, 2000; Cull, Shaughnessy, & Zechmeister, 1996; Karpicke & Roediger III, 2007b)。但是对于递增时间间隔和相等时间间隔哪一种更有优势,目前尚无一致的结论。

Landauer和Bjork(1978)最先比较了这两种不同时间间隔对测试的影响。他们让被试记忆人的姓名,并在两种条件下都进行了3次测试,但3次测试之间的间隔时间不同:递增间隔时间为1—4—10,相等间隔时间为5—5—5。学习结束后30分钟进行最终测试,发现递增间隔时间记忆成绩比相等间隔时间更好。Bjork的研究结果在当时得到了广泛的认同,研究者们认为递增间隔的提取有很大的优势,主要解释为:及时进行初次测试可以使所学的内容被成功地提取;接下来的测试间隔逐渐加长,使得每一次提取保持一定难度。Landauer和Bjork之后,研究者们继续对此问题进行了验证,但是得出的结论相差甚大。只有一部分研究结果表明递增时间间隔下的测试效应更有效(Cull, et al., 1996),有些研究认为这两种间隔类型下的测试效应没有显著差别(Balota, Duchek, Sergent-Marshall, & Roediger III, 2006; Carpenter & DeLosh, 2005; Cull, 2000),还有不少研究得出相等时间间隔下的测试的效果更好(Cull, 2000; Karpicke & Roediger III, 2007b)。Karpicke和Roediger(2007b)就发现当最终测试在10分钟后,递增间隔提取组比相等间隔提取组表现稍好,而当最终测试在2天后时,则是相等间隔提取组明显好于递增间隔提取组。

那么Bjork与Karpicke等人的观点,到底哪一个更合理呢?研究者们发现测试效应中存在一个延迟效应:在保证初次提取成功的前提下,适当延长学习与初次测试的间隔时间,增加初次提取的难度,能使学习内容保持得更久(Modigliani, 1976; Whitten, & Bjork, 1977)。依据这个现象,相等时间间隔下初次测试的延迟使得提取难度更大,从而更有利于知识的长期保持,因此虽然10分钟后在遗忘较少时递增间隔提取的表现稍好,2天后却是相等间隔提取显得更有优势。然而Bjork的实验只考察了30分钟后递增间隔提取与相等间隔提取的差异(与Karpicke的实验中10分钟后的最终测试结果一致),无法得知间隔2天后的具体情况,这样仅凭短时间间隔后最终测试上递增提取的优势就认为递增间隔提取比相等间隔提取更有效有一定的片面性。此外,Karpicke(2007b)还发现无论是在相等间隔提取组还是递增间隔提取组,被试的反应时随着测试次数增加越来越快,表明多次测试使得提取变得越来越容易,因此递增间隔提取并不符合适度学习难度原理。从上述两点,我们认为相等时间间隔的提取更有利于学习的长期保持。

不过,Karpicke(2007b)认为重要的也许不是多次测试之间的间隔如何安排,延迟初次测试才是决定两种间隔方式不同的效果的主要因素。为了验证这一假设,他在前面实验的基础上增加了一个新的变量:初次测试的间隔时间(即刻初次测试,延迟初次测试),从而产生了四种不同的测试安排:即刻初次测试后进行递增间隔测试(0—1—5—9)或相等间隔测试(0—5—5—5),延迟初次测试后继续递增间隔测试(5—1—5—9)或相等间隔测试(5—5—5—5)。该实验表明不论是在递增时间间隔还是在相等时间间隔下,延迟初次测试组两天后的记忆好于即刻初次测试组;而递增时间间隔和相等时间间隔之间并无显著差异。虽然这一结论还需要更多研究去验证,但它指出了延迟测试的重要性,这对我们应用多次测试有很大的启发。

(三)测试的形式

测试可以由多种形式,从记忆心理学角度可以包括再认、线索回忆和自由回忆等,在具体教学测试中又体现为不同的题型,如正误题、选择题、填空题、简答题、论述题等。既然测试能促进学习的长

期保持,那么不同形式的测试是否会产生不同的效果呢? Carpenter 和 DeLosh(2006)让被试记忆词语,初次测试和最终测试都采用三种形式:再认、线索回忆、自由回忆。结果得出不管最终测试采用哪种形式,初次测试为自由回忆的被试在最终测试的表现最好;初次测试为线索回忆的比再认条件下的被试在三种形式的最终测试上的表现要好。Butler 和 Roediger(2007)在更接近真实学习的情境下研究了这个问题。他们让大学生被试学习一篇文章,初次测试采用多项选择题(Multiple-choice,简称 MC)和简答题(short-answer,简称 SA),一个月后所有被试接受简答形式的最终测试。研究发现,初次测试为简答题条件下的被试在最终测试上表现较好。其他研究中也得出了相似的结论(Glover, 1989; Roediger III, Agarwal, Kang, & Marsh, 2010)。研究发现,像选择题这样的题型,其题目中直接包含正确答案,被试可依赖的线索很强,主动性较差,因此此时测试效应并不明显;此外,选择题除正确答案以外的其他选项中包含有错误信息,这很容易误导被试,形成错误记忆,在没有反馈的情况下这些错误记忆将持续存在着(Butler, et al., 2007; Roediger III & Marsh, 2005)。

需要注意的一点是,虽然题型难度越大越有利于知识的长期记忆,但它有一个重要的前提:初次测试时提取应当是成功的或者在提取失败之后提供反馈以保证成功学习。Agarwal, Karpicke 和 Kang 等人(2008)比较了重学、开卷测试、闭卷测试及有反馈的闭卷测试几种条件下学习效果的差异。从最终测试的结果看,一次闭卷测试的效果相当于重学三次(0.55vs0.54);开卷测试和有反馈的闭卷测试的成绩一样(0.66vs0.66)。可见在测试中提供反馈是非常有必要的,在初次提取失败的情况下,反馈保证了一定程度的成功学习,这样测试效应的效果才能显现出来。

(四)反馈

从日常学习经验以及前文的叙述中,可以看出反馈对利用测试促进学习的重要性是不言而喻的。虽然没有反馈的测试也能够促进被测试内容的长期保持,但提供反馈能够纠正错误从而使测试发挥更好的效果(Agarwal, et al., 2008; Butler & Roediger, 2008; Karpicke & Roediger III, 2007a),通过反馈,既利用了测试的直接效应,也发挥了测试

的间接效应。研究者们对不同类型的反馈和何时反馈效果最好进行了研究。

对测试反馈的研究发现,不同类型的反馈的确实会产生不同记忆效果。正误反馈只告诉学习者正确还是错误,而正确答案反馈给出完整的信息,这比正误反馈提供了更多的信息(Pashler, Cepeda, Wixted, & Rohrer, 2005)。除大多实验中采用的直接呈现正确答案的反馈方式外,有研究者还尝试使用了一种不断回答直至正确反馈(answer-until-correct feedback,即 AUC 反馈)的方式,即让被试重复作答直至正确为止,这种反馈需要被试主动生成正确答案,付出更多的努力,因此更有利于学习的保持。Bangert-Drowns 和 Kulik 等人(1991)对 40 个有关反馈的研究进行了元分析,其中关于反馈类型的结果表明正误反馈对提高学习的帮助很小,AUC 反馈比直接呈现完整答案对学习的长期保持更有帮助。

另一个与反馈有关的关键问题是提供反馈的时间。反馈的时间间隔主要包括即时反馈和延迟反馈。与研究者们对各种反馈类型的效果基本达成一致的观点相比,对反馈的时间间隔目前仍存在不小的争议。争议的主要原因在于目前即时反馈和延迟反馈没有明确的操作定义。一般情况下人们认为即时反馈指做完每一道题或者完成整个测试立即提供反馈,而延迟反馈中延迟多长时间则非常不确定,有的研究延迟至每道题或者每个项目做完后几秒钟,有的延迟到两天、一周甚至一个月。受斯金纳行为强化理论的影响,一些人把反馈当作是一种强化,认为出现错误后应该立即提供正确反应以消除错误反应。这个观点得到了一些研究的印证(Butler & Roediger, 2008)。然而许多研究者指出反馈与强化不是完全等同的,延迟反馈能提供充分的时间和机会让错误反应渐渐消失,从而使得在反馈后可以更容易地学习正确信息。许多研究结果也证实了这一说法(Butler, et al., 2007; Carpenter & Vul, 2011)。延迟反馈使得两次学习之间的间隔时间较长,从而对正确反应的保持更有利,这也符合学习的间隔效应说。

Kulik 和 Kulik(1988)对反馈研究的元分析发现大多数实验研究中得出延迟反馈更有效,但在实际教学研究中通常是即刻反馈更有效。Butler, Roediger 和 Karpicke(2007)认为出现这种现象的

一个最有可能的原因就是,在实验室环境下,被试会根据主试的指导语认真阅读反馈,而在实际教学中,间隔一段时间后学生的学习动机会减弱,这时他们往往快速地浏览正确答案并且只关注答错的题目,对延迟反馈的加工不够深。

三、测试促进学习相关实验研究的教育启示

从某种意义上说,教育的最终目标就是为了长时提取(包括各种知识与各种技巧),在当前教学实践中,需要记忆提取的测试得到了广泛的应用,但更多地体现在其间接作用上,特别是用测试来评价、考核学习状况,而测试本身对学习的促进作用并没有受到重视(Roediger III & Karpicke, 2006a; Roediger III, Putnam, & Smith, 2011)。随着研究的深入,将测试效应应用到教育实践和工作学习中逐渐成为一种趋势。目前已有数位研究者正开展针对测试效应应用的重要科研项目,如 Henry Roediger 正在进行的项目 TELC (Test-Enhanced Learning in the Classroom), 致力于开发出一套帮助中学教师在教学中应用测试效应的方案,使他们在应用课堂小测验时更有效、更易操作和更有灵活性。为促进我国测试效应在教育中的应用,我们应该积极吸收先前研究的有益结论,并从中受到启发:

(一) 教师要合理利用课堂测验

传统的教学环节主要包括复习提问、讲授新课、总结巩固、布置作业。像这些基本环节一样,教师也可以将小测验(quiz)作为每节课的一个基本环节,每次预留几分钟的时间来专门让学生完成小测验。根据延迟测试效应,学习后立即测试不利于知识的长期保持,因此对新学的知识点可适当延迟初次测试。初次测试后应安排多次测试以复习、巩固该知识点。Rawson 和 Dunlosky (2011) 的研究就提出了一个 3+3 的测试程序,即初次测试达到 3 次成功的提取,然后再进行 2 次复习:分别在 2 天后、1 周后。受到课堂时间和教学安排的限制,这个程序实施起来相对困难,不过教师可以在有限的时间内尽可能多地增加测试的次数。以一个单元或章节四课为例,教师可以在讲授第 2 课前进行第 1 课的测试,在第 3 课上课前再重点测试第 2 课,同时也包含少数第 1 课的核心知识,以此类推直至整个章节学习结束。在编制小测验的题目时,应以填空、简答为主,适当增加难度,以让学生自己

更主动更努力地提取。此外测试之后应当提供详细的反馈,不仅告诉学生答案是什么,也应给出深入的分析。反馈最好要延迟一段时间,教师可在每次测验后将学生试卷收齐并复印,在下次课上将详细的答案和复印的试卷一起发给学生,让学生给自己评分,然后再上交。这样既能延迟反馈,也能调动学生的积极性,从而对反馈有更深的加工。

(二) 培养学生学会和使用自我测试的学习策略

利用测试促进学习不仅是一种教学方式,更是一种学习策略,充分发挥测试对学习的促进作用更多需要依赖于学生自身。但研究发现,在大部分学生的意识里,测试只是一种评估手段而不是学习策略,因此想要让学生应用测试促进学习首要的就是要改变他们对测试作用的认识。Karpicke, Butler 和 Roediger (2009) 让大学生被试列出他们常用的学习方法,只有 11% 的学生会使用测试来帮助自己记忆;当让学生从自我测试(self-testing)和重学中选择一种来应对考试时,仅有 18% 的学生愿意选择自我测试,而且其中一半人是为了看看自己对知识的掌握情况而不是把提取当作一种学习策略。很多学生之所以会选择重学而不是测试的另一个重要原因就是测试中的提取比重学难度要大得多,学生在重学时往往自我感觉良好而在测试后会感觉自己并没有掌握多少,但是这种感觉往往是不客观的。正如 Bjork 和 Bjork (2011) 所说,“我们会被主观的感觉误导。比如重新阅读学过的章节,我们会产生一种熟悉感和知觉流畅性(perceptual fluency),于是我们觉得自己对它的理解更深了,但实际上这可能只是低水平的知觉启动。”因此,要让学生转变对学习困难的认知,不要害怕会增加困难的方法,而应当有意识地适当增加学习的难度,培养学生自我测试的学习策略。比如每学习新知识后,学生要有意识地要求自己回忆重要知识点,可以将重要概念写在卡片上,将对概念的解释写在背面,利用卡片来作线索回忆;也可学习后合上书本,将主要内容按照一定的线索梳理一遍。

(三) 设计教材与开发各类学习资源时要充分利用测试效应研究的发现

学生主要通过教材、辅导书中的习题来巩固知识点。那么对于题目的编制者来说,如何将测试效应的研究运用到编制习题中去将是个非常有价值的问题。根据前文相关研究结论,首先应当多使用

填空题、简答题,尽可能少地使用选择题;其次题目应当有一定的难度,变化具体的情境,以促进知识点的迁移,而不是简单地再现原理;新一课的习题中最好包含上一课或前几课的重要知识点,以实现重复测试;所有习题都应当提供答案及详细的解释。但使用传统的教辅用书并不能充分发挥测试效应,它在安排重复测试、延迟反馈等方面不可避免受到许多限制。目前随着计算机应用的家庭化、个人化、普及化,利用因特网实施测试或者开发出

一套基于测试效应的多媒体学习软件会更方便、快捷、有效。Rawson 和 Dunlosky(2011)的实验使用基于测试效应的提取—监控—反馈的学习方法,并且整个学习过程在计算机上完成。在被试提取之后,计算机呈现每一个概念的几个关键词和被试自己的回答,然后让被试自己给自己评分,计算机将把被试提交的结果与系统自动评分的结果进行比较,只有二者的评定一致时才能继续测试,这就使得被试必须对反馈进行充分的加工。

【参考文献】

- Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger III, H. L., & McDermott, K. B. (2008). Examining the testing effect with open and closed book tests. *Applied Cognitive Psychology*, 22(7), 861—876.
- Allen, G. A., Mahler, W. A., & Estes, W. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning and Verbal Behavior*, 8(4), 463—470.
- Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D., & Roediger III, H. L. (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology and Aging*, 21(1), 19—31.
- Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213—238.
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185—205). 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.
- Butler, A. C., Karpicke, J. D., & Roediger III, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, 13(4), 273—281.
- Butler, A. C., & Roediger, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, 19(4—5), 514—527.
- Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition*, 36(3), 604—616.
- 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. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, 19(5), 619—636.
- 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(2), 268—276.
- Carpenter, S. K., & Vul, E. (2011). Delaying feedback by three seconds benefits retention of face name pairs: the role of active anticipatory processing. *Memory & Cognition*, 39(7), 1211—1221.
- Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology*, 14(3), 215—235.
- Cull, W. L., Shaughnessy, J. J., & Zechmeister, E. B. (1996). Expanding understanding of the expanding-pattern-of-retrieval mnemonic: Toward confidence in applicability. *Journal of Experimental Psychology: Applied*, 2(4), 365—378.
- Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, 6(40).
- Glover, J. A. (1989). The “testing” phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology*, 81(3), 392—399.
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772.
- Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own. *Memory*, 17(4), 471—479.

- Karpicke, J. D., & Roediger III, H. L. (2007a). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, 57(2), 151–162.
- Karpicke, J. D., & Roediger III, H. L. (2007b). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 704–719.
- Karpicke, J. D., & Roediger III, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966.
- Kulik, J. A., & Kulik, C. L. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, 58(1), 79–97.
- Landauer, T., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. *Practical Aspects of Memory*, 1, 625–632.
- Modigliani, V. (1976). Effects on a later recall by delaying initial recall. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5), 609–622.
- Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When Does Feedback Facilitate Learning of Words? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(1), 3–8.
- Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, 330(6002), 335.
- Rawson, K., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General*, 140(3), 283–302.
- Roediger III, H. L., Agarwal, P. K., Kang, S. H. K., & Marsh, E. J. (2010). Benefits of testing memory: Best practices and boundary conditions. In G. M. Davies & D. B. Wright (Eds.), *New frontiers in applied memory*, 13–49.
- Roediger III, H. L., & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181–210.
- Roediger III, H. L., & Karpicke, J. D. (2006b). Test-enhanced learning. *Psychological Science*, 17(3), 249–255.
- Roediger III, H. L., & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(5), 1155–1159.
- Roediger III, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten Benefits of Testing and Their Applications to Educational Practice. In Mestre & Rass (Eds.), *Psychology of Learning and Motivation*, 1–36.
- Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology*, 30(9), 641–656.
- Tulving, E. (1967). The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 6(2), 175–184.
- Wheeler, M. A., & Roediger, H. L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science*, 3(4), 240–245.
- Whitten, II., & Bjork, R. A. (1977). Learning from tests: Effects of spacing. *Journal of Verbal Learning and Verbal Behavior*, 16(4), 465–478.

(责任编辑 侯珂 责任校对 侯珂 蒋重跃)

Testing for Learning Enhancement: Recent Advances in Memory Psychology and Educational Revelation

LUO Liang, ZHANG Wei

(National Key Laboratory of Cognitive Neuroscience and Learning, BNU, Beijing 100875, China)

Abstract: Memory Psychology researches reveal that testing can not only be a means of assessing the previous learning effect, but also change the information in Long-term Memory. Compared with simple repetitive learning, one or more than one testing in learning can apparently enhance the maintenance of the learnt content in Long-term Memory. This enhancement has gained manifold empirical supports and the researchers have offered different theoretical interpretations, including memory disuses based on theoretical mechanism and elaborative retrieval hypothesis for emphasizing the internal apparatus. The factors that may influence on learning enhancement by testing include testing times, duration, form, and feedback. Currently, the role of testing on enhancing learning itself should be noted. Then teachers are required to make use of classroom testing, and to foster students' learning strategies for using self-testing, and the relevant research findings.

Key words: testing effect; memory retrieval; study; memory psychology