

# 测验对学习促进作用的实验研究及启示

李宏英<sup>1</sup>, 张洁尉<sup>2</sup>, 连榕<sup>1</sup>

(1. 福建师范大学教育科学与技术学院, 福建福州 350007; 2. 西南大学心理学院, 重庆 400715)

**摘要:** 考察了学习中的测验对最终测验回忆正确率的影响。被试内因素是学习过程(重复学习、有测验的学习), 被试间因素为学习与最终测验之间的时间间隔(10分钟、1天)。发现有测验的学习条件优于重复学习条件, 即存在测验效应, 间隔10分钟组的正确率显著高于间隔1天组。测验可以作为学习的一种方式, 在教学中得到更加合理的使用。测验与教学相结合的策略对于基础教育改革实践是有益的尝试。

**关键词:** 重复学习; 测验效应; 教学策略; 基础教育改革

**中图分类号:** G442      **文献标识码:** A      **文章编号:** 1008-0627(2009)05-0032-04

## 一、问题的提出

教育与心理学研究者一直在不断探索教育实践中的不同因素对学生学习效果的影响。测验是教育实践中的一个重要环节, 教学中教师对学生进行测验的目的通常是检验学生对已学知识的掌握程度。但在心理学原理指导下, 在学习过程中有目的地加入测验还可以帮助学生巩固已学知识并提高学习效果。这就是1967年Tulving所提出的“测验效应”(testing effect), 也称为“考试效应”,<sup>[1]</sup>即测验能促进学生在后续阶段中回忆已学的知识。学者们从不同的角度验证了测验效应, 例如Hogan和Kintsch让被试学习一个词表四次(SSSS)或者学习一次后回忆三次(STTT), 并于两天后进行总的自由回忆测验, 考察了不同的学习过程的影响, 结果发现有测验的学习过程比学习四次的学习过程的记忆效果要好。<sup>[2]</sup>还有研究发现, 如果在学习后立即进行一个测验, 即使在学习过程的测验中没有关于正确或错误的反馈, 并且也没有进行更深入的学习, 那么在一段时间后对该学习内容的测验中, 也会产生更好的成绩。<sup>[3]</sup>这方面的研究总体上来看不仅有实验室研究, 也有实际教学背景下的研究; 还使用了词表、配对词及散文等不同的实验材料以及不同的测验形式, 如单选、填空、判断等客观题型和简答等主观题等等。为便于实验控制, 关于测试效应的实验室研究大多采用词表或无关词对等人工材料, 但真实教育情境中的学习材料与词表或无关词对毕竟是不同的。也有研究采用了科学小品文等真实的教育素材做为实验材料<sup>[4]</sup>, 但这类研究大多采用了被试间设计。

基于以往的研究情况, 当前的实验采用了教育实践中的学习材料做为实验材料, 并在学习过程变量上采用了被试内设计以提高实验的精度。实验目的是考察学生在包含一个测验的学习过程中和仅仅重复学习(相当于通常的复习)的学习过程中的学习效果, 并考察在不同的时间间隔(10分钟、1天)后不同学习过程中所学材料的掌握程度是否存在差异。实验预期存在测验效应, 表现为相比之下有测验的学习过程比重复学习的学习过程具有更好的学习效果, 时间间隔与学习过程存在交互作用, 即不同的学习过程(重学或测验)在不同的时间上表现出不同的效果。

## 二、方法

### (一) 被试

某大学学生24人, 平均年龄18岁, 男女性别各半, 视力或矫正视力正常, 此前未参加过类似的实验。实验得到有效数据21份。

### (二) 实验设计、材料和仪器

实验采用二因素混合设计: 一个因素是学习过程变量, 分为测验和重学两个水平, 此变量为被

试内变量。另一个因素是学习阶段与最终测验阶段之间的时间间隔变量，包括10分钟和1天两个水平，此变量为被试间变量。因变量为被试在最后的有线索回忆测验中的正确率。

实验材料为4篇节选自托福试题(1991~1992)的知识性短文。4篇短文的主题分别为雨水(431字)、太阳(538字)、信鸽(451字)和海狸(469字)。学习阶段进行的测验为选择测验，每篇短文对应3个选择题(4选1)，共12个问题，要求被试从4个备选答案中选择正确的一项(其中1个是正确词项，其他3个是诱惑词项)。最终测验阶段的测验是以填空形式进行的有线索提示测验(cued-recall test)。短文的学习和测验均在分辨率为1024×768的DELL微机上进行。文章的文字为黑色宋体小三号字，程序使用Eprime1.1编制，使用SPSS15.0对数据进行分析。

### (三) 实验程序

实验包含三个阶段：学习阶段、填充任务阶段和最终测验阶段。学习阶段：先学习1篇短文(60秒)，然后重学短文一遍或根据短文内容完成相应的选择测验(60秒)，随后学习下一篇短文。每学习完一篇休息10秒，依次学完4篇短文。为了平衡顺序效应，本实验采用了ABBA法，一半被试按照“重学—测验—测验—重学”的顺序进行，另一半被试按照“测验—重学—重学—测验”的顺序进行。填充任务阶段：在学习阶段结束后，要求被试完成屏幕上出现的数学运算，例如： $3+43+86=?$ ，此阶段持续进行10分钟。最终测验阶段：测验为以纸笔的填空形式进行的线索回忆测验。在填空测验结束后，随机将被试分为人数相等的两组，对其中一组被试立即进行测验(10分钟组)，要求另一组被试于1天后回来完成一个小任务，对1天组并没有提到1天后的小任务就是测验。问卷里要求被试填写编号、性别、年龄等基本信息。

### 三、结果与分析

最终测验阶段不同学习过程和时间间隔下的正确率见表1。关于学习过程(重学、测验)与测验时间间隔(10分钟、1天)的重复测量一个因素的二因素方差结果见表2。

从方差分析结果表明，测试时间的主效应显著， $F(1, 19)=20.58$ ， $p<0.01$ ，无论有无测验，10分钟组比1天组的正确率更高；学习过程的主效应也显著 $F(1, 19)=34.06$ ， $p<0.01$ ，最终测验时，无论是10分钟组还是1天组，测验条件都表现出更高的正确率。但是测验时间与学习过程的交互作用不太显著 $F(1, 19)=0.56$ ， $p>0.05$ 。

在测验学习条件下，被试的最终填空测验的错误来源基本上有两类，以测验中的诱惑词来填空的错误(诱惑词错误)和以非诱惑词填空的错误(非诱惑词错误)，两种来源的错误率情况如表3所示。关于错误来源(诱惑词、非诱惑词)与测验时间间隔(10分钟、1天)的重复测量一个因素的二因素方差结果见表4。统计分析结果表明，错误来源的主效应不显著，“非诱惑词错误率”与“诱惑词错误率”二者差异不显著， $F(1, 19)=1.12$ ， $p>0.30$ 。测验的时间效应显著， $F(1, 19)=16.27$ ， $p<0.01$ 。时间与错误来源交互作用不

表1 不同学习过程和时间间隔后最终测验正确率(M±SD)

时 间	重 学	测 验
10分钟	0.38±0.18	0.72±0.19
一天	0.17±0.18	0.42±0.14

表2 不同学习过程在不同测验时间中正确率的方差分析

项 目	df	MS	F	p
时间间隔	1	0.68	20.58***	0.00
学习过程	1	0.93	34.06***	0.00
时间间隔 X 学 习过程	1	0.02	0.56	0.46
误差	19	0.03		
总计	22			

注：p<0.05\*，p<0.01\*\*，p<0.001\*\*\*。

表3 测验学习条件中诱惑词和非诱惑词的错误率(M±SD)

时 间	诱惑词错误率	非诱惑词错误率
10分钟	0.10±0.18	0.18±0.17
一天	0.24±0.20	0.33±0.26

注：诱惑词错误率、非诱惑词错误率，以及表1中的测验正确率三者相加之和为1。

表4 两种测验时间条件对应的最终测验错误率方差分析

项 目	df	MS	F	p
错误来源	1	0.08	1.12	0.30
时间间隔	1	0.11	16.27***	0.00
时间间隔 X 错误来源	1	0.00	0.01	0.96
误差	19	0.07		
总计	22			

注：p<0.05\*，p<0.01\*\*，p<0.001\*\*\*。

显著,  $F(1, 19)=0.01, p > 0.96$ 。

#### 四、讨论

##### (一) 关于测验效应的讨论

实验中, 重复学习条件下学习短文 2 遍, 测验条件下学习短文一遍, 但无论时间间隔为 10 分钟还是 1 天, 最终测验时测验条件下所学习材料的正确率均高于重学条件, 表现出了“测验效应”, 这与以往的研究结果是一致的。<sup>[5]</sup>

该实验结果可以用适当转换理论 (theories of transfer-appropriate processing, TAP 理论)<sup>[6]</sup>来解释, 该理论强调学习和测验过程所需要操作的一致性, 学习过程中的测验会对最后测验的效果有积极的促进作用, 是因为最终测验过程所需要的知识提取技巧, 无论是外显的还是内隐的技巧均已经在学习过程的测验中得到了练习, 而重复学习的过程则不能提供这种练习的机会。

实验并未发现学习过程 (重学、包含测验) 与测验时间的交互作用, 10 分钟与 1 天这两个测验时间里, 均表现为测验条件下的学习优于重学条件下的学习。该结果与艾宾浩斯提出的记忆的遗忘规律相一致, 但与前人的研究并不一致, 例如 Roediger 和 Karpicke<sup>[4]</sup>的实验中表现出了测验时间与学习过程的显著的交互作用, 即学习后进行的即时测验中, 重学条件的记忆效果优于测验条件, 延时测验中, 测验条件的效果优于重学条件。可以从两个方面分析这种实验结果差异:

首先, 这种差异与本次实验所使用的短文的数量、短文所附带题目数量有关。前人的研究中的实验材料以词表为主, 以短文作为实验材料的研究中短文篇数一般较多 (如 12 篇), 短文附带题目数量一般为每篇 7 题甚至更多。当前测验中所使用的文章所带的题目均为 3 题, 这从一定程度上限制了被试的作答。

其次, 在重学条件中, 短文的其他内容对最终测验的知识点具有一定的干扰作用; 在测验条件中, 由于所有的题干以及正确选项均与最终测验的考察知识点有关, 被试的学习目标更为具体和明确。由此, 无论在 10 分钟还是 1 天以后的测验中, 都表现出了测验条件比重学条件更好的回忆率。

##### (二) 关于消极指导效应的讨论

从统计分析结果来看, 错误来源的主效应不显著, 即“非诱惑词错误率”与“诱惑词错误率”二者差异不显著, 因此并未显示出“消极指导效应” (the negative suggestion effect) 现象, 即当学习中插入测验时, 存在“测验效应”的同时, 也存在着增强学生们从测验的诱惑词项中得到错误信息的可能。一些研究曾经证实了消极指导效应, 例如, Jacoby 与 Hollingshead 的研究表明, 呈现给被试错误拼写的单词会增加口头拼写测验的错误率。<sup>[7]</sup>在 4 选 1 的选择题中, 只有一个选项是正确的, 如果被试认真看了每个选项再做出选择, 那么就相当于他们要看 3 个错误选项, 再看一个正确选项。这些错误选项就可能成为下次测验时的干扰, 从而对测验产生消极的影响。但本研究结果并未显示出这种测验带来的消极指导效应, 这可能与本实验中题目数少有关, 这也是本研究需要改进的方面, 在今后的研究中应适当增加所选文章的篇数及文章所带的题数, 今后关于消极指导效应还要做进一步的探讨。

##### (三) 对教育实践的启示

提高课堂教学效果是基础教育改革实践中的一个重要环节。2001 年 6 月颁布的《基础教育课程改革纲要》指出: “教师在教学过程中应与学生积极互动、共同发展, 要处理好传授知识与培养能力的关系, 注重培养学生的独立性和自主性, 引导学生质疑、调查、探究, 在实践中学习, 促进学生在教师指导下主动地、富有个性地学习……创设能引导学生主动参与的教育环境, 激发学生的学习积极性, 培养学生掌握和运用知识的态度和能力, 使每个学生都能得到充分的发展。”<sup>[8]</sup>

测验效应的研究结果对教育实践具有重要的启示。教师在教学过程中可以采用测验与教学相结合的教学策略, 即在讲授主要教学内容之后立即进行测验。这种测验与教学相结合的教学策略对于教学效果具有积极的影响。首先, 本研究结果以及前人的相关研究结果均证实, 与反复复习相比, 在学习过程之后立即进行测验可以取得更好的记忆效果。其次, 研究结果证明, 在学习过程中进行测验的学生, 不仅记忆程度在一段时间之后保持较好的水平, 而且对学习材料的学习兴趣更高, 对

自己的学习效果具有充足的信心。因此，通过对课堂测验的有效运用，不仅能够有效地增强学习效果，而且可以提高对学习的兴趣，进而提高学习的主动性。更为重要的是这种教学策略是一种实现了学生整体参与的教学策略，所有学生均可以在这个学习——测验环节中受益，避免了采用课堂提问或讨论等其它策略时学生只能部分参与的问题。因此测验与教学相结合的策略在教育实践中具有极大的教育价值，对于基础教育改革实践而言将是一个有益的尝试。

### 五、结论

实验对“测验效应”进行了研究，实验发现：

1. 实验结果证明存在“测验效应”。学习阶段有测验的学习过程比仅仅重复学习的学习过程使被试表现出更高的回忆正确率，即测验能够增强学习效果，有测验的学习效果优于无测验的学习。

2. 测验效应会随着测验时间的推移而逐渐降低。综上所述，有测验的学习效果优于单纯的复习。测验不仅可以衡量学生对知识的掌握程度，还可以对学习起到促进的作用。测验效应研究结果为在教学中使用测验提供了更多空间，测验可以作为学习的一种方式，在教学中得到更加合理、充分的使用，对于基础教育改革具有一定的教育价值。但测验的积极作用还与所使用的测验题目以及其他一些因素有关，今后的研究中还可以从选项数、测验的反馈、测验的形式等方面入手，对测验效应进行更加深入的研究，进而得到更加丰富的并有意义的结论。

### 参考文献

- [1] TULVING E. The effects of presentation and recall of material in free-recall learning[J]. *Journal of Verbal Learning and Verbal Behavior*, 1967(6): 175-184.
- [2] HOGAN R M, KINTSCH W. Differential effects of study and test trails on long-term recognition and recall[J]. *Journal of Verbal Learning and Verbal Behavior*, 1971(10): 562-567.
- [3] RONALD J, NUNGESTER. Testing versus review: effects on retention[J]. *Journal of Education Psychology*, 1982, 74(1): 18-22.
- [4] ROEDIGER H L, KARPICKE J D. Test enhanced learning: Taking memory tests improves long-term retention[J]. *Psychological Science*, 2006, 17(3): 249-255.
- [5] KARPICKE J D, ROEDIGER H L. The critical importance of retrieval for learning[J]. *Science*, 2008(319): 966-968.
- [6] ROEDIGER H L, GALLO D A. Processing approaches to cognition: the impetus from the levels-of-processing framework [J]. *Memory*, 2002(10): 319-332.
- [7] JACOBY L L, HOLLINGSHEAD A. Reading student essays may be hazardous to your spelling: effect of reading incorrectly and correctly spelled words[J]. *Canadian Journal of Psychology*, 1990(44): 345-358.
- [8] 中华人民共和国教育部. 基础教育课程改革纲要(试行)[N]. 中国教育报, 2001-07-27(02).

## An Experimental and Suggestive Study of the Stimulus Effects of Testing on Learning

LI Hong-ying<sup>1</sup>, ZHANG Jie-wei<sup>2</sup>, LIAN Rong<sup>1</sup>

(1. College of Educational Science and Technology, Fujian Normal University, Fuzhou 350007, China; 2. School of Psychology, Southwest University, Chongqing 400715, China)

**Abstract:** The purpose of the paper is to discuss the effects of formulative tests on the rate of retention of summative tests in terms of subjects' learning style (review, test) and interval (10min, 1day) between study and the final test. The writer finds that the testing effects of test-based study are superior to those of review. A significant difference in the proportion of correct responses stands significantly between 10min interval and 1 day interval. Testing, an aid to learning, can be adopted in teaching in an effort to promote the reform of basic education.

**Key words:** testing review; testing effect; teaching strategies; reform of basic education

(责任编辑 赵 蔚)