

# 课程嵌入型表现性评定对数学问题解决影响的实验研究<sup>\*</sup>

张咏梅<sup>1</sup> 刘红云<sup>2</sup> 孟庆茂<sup>2</sup>

(1. 北京教育科学研究院, 北京 100031; 2. 北京师范大学心理学院, 北京 100875)

**摘 要:**课程嵌入型表现性评定是当前基础教育新课程改革倡导的重要的学生学业评价方式。以 335 名初中一年级学生为被试, 采用为期一学年的追踪实验研究方式, 探索课程嵌入型表现性评定对数学问题解决的影响效果。研究结果表明, 此方式对学生的数学问题解决能力发展具有显著促进作用, 且促进作用随实验进程不断增加并受到学生原有学业水平的影响。对口语报告资料的分析进一步表明, 此评定方式对数学问题解决的促进作用主要体现在理解问题、元认知与问题解决策略等方面。

**关键词:**表现性评定; 课程嵌入型表现性评定; 数学问题解决

## 1 问题提出

当前, 我国基础教育课程改革处于全面深入推进过程。表现性评定, 作为新课程改革所倡导的学生学业评价方式已引起了极大的关注, 并越来越广泛地被应用到教学实践中去。广义上来讲, 表现性评定是指与多项选择测验相对应的一切题目形式, 从狭义上来讲是指为测量学习者运用先前所获得的知识解决新异问题或完成特定任务能力的一系列尝试, 即运用真实的生活或模拟的评价练习来引发最初的反应, 由高水平评定者按照一定标准进行直接的观察、评判, 形式包括建构式反应题、书面报告、作文、演说、操作、实验、资料收集、作品展示<sup>[1]</sup>等。

对表现性评定的科学性研究开始于上世纪 80 年代后期的国外教育评价与测量领域, 其兴起具有深刻的认知理论、建构主义理论与人本主义理论的背景。依据操作过程风险性与限制性程度, 表现性评定可以划分为课程嵌入型表现性评定与大规模测验中的表现性评定。前者指以教师为评价主体, 以学生日常学业表现为评价客体的表现性评价形式, 特点在于与所学课程内容匹配, 风险性较低, 属于校内评价范畴<sup>[2]</sup>。课程嵌入型表现性评定是本文研究的焦点(以下统称表现性评定)。

Brenda 在研究中指出, 测量专家对表现性评定的研究不应只强调任务的表面特征, 对学生兴趣的

调动, 还应该考虑到与以问题解决能力为代表的认知概念(如问题解决等)的关系及其影响<sup>[3]</sup>。二十多年来, 有许多教育类的理论研究支持表现性评定在日常教学背景中的运用, 并明确提出表现性评定对问题解决等高级思维能力的发展具有促进作用<sup>[4]</sup>。1999 年, Fuchs 等采用情景实验研究的方法发现, 与非表现性评定条件下的学生相比, 表现性评定条件对不同学业水平学生的数学问题解决能力的发展均有促进作用<sup>[5]</sup>。

表现性评定对数学问题解决能力的影响正在引起越来越多的关注。国内外许多教育理论研究对表现性评定的实施效果推崇同时, 大体还基本停留在理论研究阶段, 规范性的实证研究极为缺乏。在已有极少的实验研究如 Fuchs 等人的研究中, 往往将如问题解决等认知能力作为研究整体来处理, 且多在情景实验中完成。而研究者认为只有实践性、追踪性与分析性的研究, 才会更好地体现表现性评定对问题解决能力的影响, 为教学提供更有价值的信息与建议。

以上述研究为基础, 本研究的目的是结合我国当前的课程改革的背景, 以表现性评定的实施作为自变量, 以数学学科为载体, 以数学问题解决能力为因变量, 采用追踪研究的方式, 在较长的时间内考查表现性评定对于数学问题解决能力及其发展变化的影响。在此, 数学问题解决被界定为“事先不知道解

<sup>\*</sup> 基金项目: 北京市教育科学研究“十五”规划青年专项课题(BJ15CGA02304)资助。

作者简介: 张咏梅(1972—), 女, 北京教育科学研究院副研究员。E-mail: zhyzm72@gmail.com

决方法的任务求解过程”<sup>[6]</sup>。1945 年, Polya 首先提出了数学问题解决的理论模型框架, 认为数学问题解决主要包括理解问题、制订解决方案、执行方案与检查等步骤; 上世纪 80 年代中期, Mayer 则认为解决数学问题要求四类知识或过程, 即转换、整合、计划与监控、实施解决方案等; 上世纪 90 年代以来, Glaser 提出问题解决的综合模型, 认为数学问题解决应包括四个部分: 具体领域知识、元认知、问题解决策略、情感和动机因素。几十年来, 由于 Polya 模型完整性及与数学问题解决过程的接近性, 一直是数学教学研究的热点; Mayer 的框架则以侧重问题解决心理机制而见长; 近年来, 随着对问题解决过程中元认知、问题解决策略与动机因素的认识不断加深, 以 Glaser 模型为代表的数学问题解决模型的影响不断扩大<sup>[7]</sup>。

根据表现性评定与数学问题解决的已有研究资料, 结合研究目的, 本研究的具体问题是表现性评定的实施是否对于数学问题解决及其发展变化具有一定的影响; 影响主要体现在数学问题解决的哪些具体过程或方面; 影响是否受到学校类型因素、性别因素与学生原有学业水平因素的作用。

## 2 研究方法

### 2.1 被试

北京市初一学生 335 名, 分别来自于市重点、区重点与普通中学。4 个实验班学生 165 人, 4 个控制班学生 170 人; 男生 188 人, 女生 147 人。

同一学校的实验班与控制班学生在初中入学后的摸底考试中成绩不存在显著性差异, 且 2003 年 1 月的期末数学考试成绩不存在显著性差异。实验班与控制班的男女生数量无显著差异。

### 2.2 方法

追踪(纵向)研究法、测验法、口语报告法。相关录音设备与转写工具。

### 2.3 编制工具

**2.3.1 编制数学表现性评定方案:** 由初中数学学科专家、数学学科教师与教育评价的研究者按照全日制义务教育数学学科新课程标准的要求, 结合华东师大版的初一、初二年级的数学教材进度设计十个表现性评定任务方案(见附件 1)。采用专家判断法, 以《数学表现性评定任务内容效度评价表》(见附件 2)为工具, 对任务方案的内容效度进行评价。19 名初中数学学科专家对表现性评定任务方案的内容效度的综合评价结果达到 6.55, 标准差为 1.60, 在

包括 0~9 的十等级的评价表中处于较高等级。单因素方差分析的结果表明, 十个任务在综合评定等级方面不具有显著差异( $F_{(9, 180)} = 0.50, p > 0.05$ )。

**2.3.2 编制数学问题解决测验:** 共包括 3 份, 主要是为测量数学问题解决能力。由 4 位数学学科专家结合初中数学课程进度编制而成, 其内部一致性  $\alpha$  系数分别为 0.66, 0.66, 0.56。满分为 80 分。

## 3 研究程序

**3.1** 在 2003 年 2 月—2003 年 12 月的两个学期中, 实验班实施表现性评定任务方案(见附件 1), 控制班实施普通作业方案。即实验班学生完成 10 个表现性评定任务, 每个约需要 2~3 周的时间, 控制班学生按照课程要求完成日常普通作业方案。表现性评定任务的实施程序主要包括以下环节:

**3.1.1 准备环节** 在整体实验开始前, 为每个学生建立数学作业档案袋, 班中建立班级展示园地;

**3.1.2 作业呈现环节** 要求在某个单元、或章节内容将近结束时呈现表现性评定任务。呈现时, 讲解表现性评定的任务要求与任务评价维度。结合实例告知学生们什么样的作品是一件好的作品, 什么样的表现是一种好的表现;

**3.1.3 作业完成环节** 要求完成作业后, 首先是在指定的时间内个人自评, 然后上交至小组讨论评定, 最后上交至教师评定。教师评定时, 评语要采用肯定、鼓励的语言风格, 明确肯定优点, 指出不足;

**3.1.4 反馈环节** 要求以小组为单位汇报本小组的作业完成情况, 结合作品本身, 指出小组作品的优势, 并就具体存在的问题进行简单阐述。最后由教师针对有代表性的问题结合教材内容进行总结。对自己完成任务情况不满意的同學, 可以将作品拿出, 重新完成, 如果在规定时间内上交, 将不影响成绩;

**3.1.5 作品展示与汇集** 要求在某表现性评定任务完成后, 将所评出的比较优秀的作品在班级园地中进行展示。并要求学生们将自己的作品放入档案袋中保存, 留待期末时进行作品集展示会, 并邀请家长来参加。

**3.2** 分别于 2003 年 3 月、7 月、12 月, 对实验班与控制班学生实施《数学问题解决测验》(见附件 3)。在 2003 年 3 月与 12 月的测验中, 《数学问题解决测验》有 A 与 B 两种版本。A 版本由 4 道大题组成, B 版本包括 A 版本中的 3 道试题。

**3.3** 在 2003 年 3 月与 12 月的测试中, 分别从实验班与控制班中各抽取 1 名数学日常表现中上与中下

的被试,共计 16 名,其中男女生各半。要求他们在单独的房间中完成 B 版本,同时,实验班与控制班中的其他同学完成 A 版本。(在 12 月的口语报告测试中,1 名学生因病缺失。)

3.4 按照口语报告操作指导,采用即时口语报告与追溯口语报告相结合的方式<sup>[8]</sup>,要求 16 名被试在单独测试过程中“大声报告出自己思考过程”。并在被试解决完问题或放弃解决问题之后,让其回忆并再次阐述自己“问题解决的过程”。在此过程中,利用设备进行录音。

3.5 选择两位具有 5 年初中数学教学经验的教师与《数学问题解决测验》的编制者一起制订、讨论题目的标准答案。对试卷进行试判,再次统一标准后,开始正式阅卷。每张试卷阅两遍,每位学生成绩取两位阅卷教师所给成绩的平均值。

3.6 建立编码系统,对口语报告资料进行编码

首先把 16 名被试的数学问题解决口语报告资料逐字逐句转写为书面材料。在多次通读所有访谈资料的基础上,以 Polya 和 Glaser 提出的问题解决理论框架为基础<sup>[7]</sup>,形成口语报告资料编码系统(见附件 4),来对数学问题解决的过程进行深入的分析。

利用该编码系统对访谈资料进行试编码,并通过试编码进一步修改编码系统。通过反复的试编码和对编码系统的修改,形成正式的编码系统,其包括 5 个子维度(一级编码)分别为理解问题、具体领域知识、元认知、问题解决策略和努力。

由受过培训的两名编码者(研究者本人和一名基础心理学专业的硕士研究生)按照正式编码系统对口语报告资料进行编码。在两次共 32 份口语报告资料中,随机抽取出 4 份。两名编码者对 4 份资料独立进行编码,其一致性为 81%。对编码过程中意见不一致的编码,通过两人讨论确定。与此同时,编码者对被试在每个一级编码上的整体效果情况在 1~5 的标尺上进行判断。其中 1 表示整体效果差,2 表示整体效果较差,3 表示整体效果中等,4 表示情况整体效果较好,5 表示整体效果好。

## 4 研究结果与分析

研究结果采用软件 SPSS10.0 与 AMOS4.0 进行分析。根据各部分研究的数据类型,主要采用重复测量多因素方差分析法、非参数检验法、路径分析法来进行分析。

4.1 表现性评定对学生数学问题解决能力(下称 MPS)的影响

表现性评定实验对于学生 MPS 发展的影响。由表 1 所示,对于本部分重复测量多因素方差分析结果表明,主效应 MPS 的三次测量与实验处理在统计上都达到显著( $F_{(2,612)}=18.85, p<0.001, F_{(1,306)}=9.56, p<0.01$ ),说明实验班与控制班之间在 MPS 上存在着显著差异;交互作用 MPS×实验达到显著水平( $F_{(2,612)}=6.20, p<0.01$ ),这说明在三次测量中,实验班与控制班差异趋势不同。如图 1 所示,在第一次测量 MPS 时,实验班略高于控制班( $39.84-38.99=0.85$ ),但不存在 MPS 上的显著差异( $t=0.48, p>0.05$ );第二次测量 MPS 时,两班差距拉大( $47.21-40.51=6.70$ ),实验班显著地高于控制班( $t=3.72, p<0.001$ );在第三次测量 MPS 时,差距基本未变( $41.72-35.12=6.60$ ),实验班还是显著高于控制班( $t=3.24, p<0.001$ )。因此,数据分析结果表明表现性评定对学生的数学问题解决能力具有促进与提高的作用。

表 1 表现性评定对学生 MPS 影响的重复测量方差分析表

变异来源	方差	自由度	均方	F
组内				
MPS	5163.50	2	2581.75	18.85***
MPS×实验	1723.15	2	861.57	6.20**
误差	83825.51	612	136.97	
组间				
实验	5136.97	1	5136.97	9.56**
误差	164508.11	306	537.61	

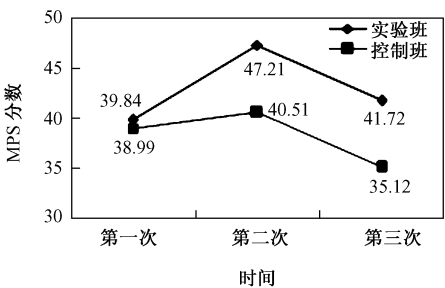


图 1 MPS×实验交互作用分析图

4.2 表现性评定对 MPS 具体领域的影响

对数学问题解决口语报告资料编码进行分析,由表 2 具体领域人均得分率可以看出,在第一次数学问题解决口语报告中,实验班在理解问题、具体领域知识、元认知、问题解决策略与努力方面的表现与控制班基本相同。其中,在理解问题整体方面,实验班略差于控制班( $0.70<0.78$ ),进一步分析表明,除重读题目外,在图表理解、表达题目与信息标记方面都略差于控制班;在具体领域知识方面,实验班好于控制班( $0.94>0.69$ ),主要体现在基本概念的运用

方面与定理与公式的运用方面;在元认知方面,两班基本相同,实验班在制定计划与监控进程方面稍好,控制班在检查改正方面稍好;在问题解决策略方面,实验班稍好于控制班( $0.88>0.69$ ),主要体现在启发式运用稍多;两者在努力方面相同。

表 2 实验班与控制班口语报告学生在  
在数学问题解决能力各领域的具体表现  
(人次(具体领域人均得分率))

	第一次		第三次	
	实验班	控制班	实验班	控制班
理解问题	23(0.70)	25(0.78)	27(0.84)	18(0.64)
重读题目	7(0.88)	6(0.75)	6(0.75)	4(0.57)
图表理解	4(0.50)	5(0.63)	6(0.75)	4(0.57)
表达题目	7(0.88)	8(1.00)	7(0.88)	5(0.71)
信息标记	5(0.63)	6(0.75)	8(1.00)	5(0.71)
具体领域知识	15(0.94)	11(0.69)	16(1.00)	13(0.81)
概念	8(1.00)	5(0.63)	8(1.00)	7(1.00)
定理与公式	7(0.88)	6(0.75)	8(1.00)	6(0.86)
元认知	20(0.66)	2(0.63)	21(0.66)	11(0.34)
认识特点	3(0.37)	3(0.37)	3(0.37)	1(0.14)
制定计划	8(1.00)	7(0.88)	7(0.88)	3(0.43)
监控进程	8(1.00)	7(0.88)	8(1.00)	6(0.86)
检查改正	2(0.25)	3(0.37)	3(0.38)	1(0.14)
问题解决策略	14(0.88)	11(0.69)	16(1.00)	10(0.72)
启发式	7(0.88)	4(0.50)	8(1.00)	4(0.57)
解决技巧	7(0.88)	7(0.88)	8(1.00)	6(0.86)
努力	5(0.63)	5(0.63)	4(0.50)	3(0.43)

在第三次数学问题口语报告中,实验班在理解问题、元认知、问题解决策略方面明显好于控制班,主要体现在重读题目、图表理解、表达题目与信息标记等各个方面。具体来说,在理解问题方面( $0.84>0.64$ ),主要体现在图表理解、表达题目与信息标记方面;在元认知方面( $0.66>0.34$ ),主要体现在认识特点、制定计划与检查改正上;在问题解决策略方面( $1.00>0.72$ ),主要体现在启发式运用增加,解决问题的技巧性加强。在具体领域知识与努力方面,两者间的差距不明显。

4.3 表现性评定对不同学校学生 MPS 发展变化的影响

由表 3 所示,重复测量多因素方差分析结果表明,主效应 MPS、实验处理、学校类型在统计上都达到显著水平( $F_{(2,604)} = 13.14, p<0.001; F_{(1,302)} = 12.68, p<0.001; F_{(2,302)} = 19.31, p<0.001$ )。学校类型主效应的检验表明,差异主要存在于市重点校与区重点校之间,市重点校与普通校之间( $47.12>40.12, 40.12>36.57$ );两次交互作用 MPS×学校达到显著水平( $F_{(4,604)} = 15.30, p<0.001$ ),说明在三次 MPS 测量中,各类学校的差别情况存在显著差

异。深入分析表明,在第一次测量中,市重点校显著高于区重点校与普通校( $50.99>37.52>33.19$ ),但在第二次测量中,学校间的 MPS 差距减小,不存在显著差异;在第三次测量中,差距又有增加的趋势,市重点校依然显著高于区重点校与普通校( $45.63>39.46>33.32$ )。

表 3 表现性评定对不同学校学生 MPS 影响的  
重复测量方差分析表

变异来源	方差	自由度	均方	F
组内				
MPS	3292.34	2	1646.17	13.14***
MPS×实验	1760.50	2	880.25	7.03**
MPS×学校	7666.16	4	1916.54	15.30***
MPS×实验×学校	549.59	4	137.40	1.10
误差	75657.70	604	125.26	
组间				
实验	6090.31	1	6090.31	12.68***
学校	18547.92	2	9273.96	19.31***
实验×学校	842.17	2	421.09	0.88
误差	145066.89	302	480.35	

与此同时,两次交互作用实验×学校未达到显著水平( $F_{(2,302)} = 0.88, p>0.05$ ),说明实验班与控制班之间的差别在各类学校间没有显著差异;三次交互作用 MPS×实验×学校也未达到显著水平( $F_{(4,604)} = 1.10, p>0.05$ ),说明三次 MPS 测量中,实验班与控制班的差别情况在各类学校间没有显著差异。

4.4 表现性评定对不同性别学生 MPS 发展变化的影响

由表 4 所示,重复测量多因素方差分析结果表明,主效应 MPS、实验处理、性别都在统计上都达到显著水平( $F_{(2,608)} = 19.33, p<0.001; F_{(1,304)} = 9.51, p<0.01; F_{(2,302)} = 4.63, p<0.01$ ),对性别主效应的检验表明,女生的 MPS 显著高于男生( $42.39>39.10$ );两次交互作用 MPS×性别达到边缘显著水平( $F_{(2,604)} = 2.87, p=0.057$ ),说明在三次 MPS 测量中性别之间的差别情况可能存在显著差异。深入分析表明,虽然在每次测量中,均为女生的 MPS 较男生高,但是三次的差异情况不同。如图 2 所示,在第一次测量中,女生与男生的 MPS 基本相同, ( $39.78>39.11$ ),但在第二次与第三次测量中,男女生间的 MPS 差距增加了,女生高于男生 ( $46.44>41.67, 40.89>36.32$ )。

与此同时,两次交互作用实验×性别未达到显著水平( $F_{(1,304)} = 0.04, p>0.05$ ),说明男女生间的差别在实验班与控制班之间没有显著差异。三次交

互作用 MPS × 实验 × 性别也未达到显著水平 ( $F_{(2,604)} = 0.44, p > 0.05$ ),说明三次 MPS 测量中,男生与女生的差别在实验班与控制班之间没有显著差异。

表 4 表现性评定对不同性别学生 MPS 影响的  
重复测量方差分析表

变异来源	方差	自由度	均方	F
组内				
MPS	5279.14	2	2639.57	19.33***
MPS×实验	1668.84	2	834.42	6.11**
MPS×性别	783.98	2	391.99	2.87
MPS×实验×性别	0.87	2	0.44	0.00
误差	83039.54	608	136.58	
组间				
实验	5068.38	1	5068.38	9.51*
性别	2468.49	1	2468.49	4.63*
实验×性别	20.40	1	20.40	0.04
误差	162026.16	304	532.98	

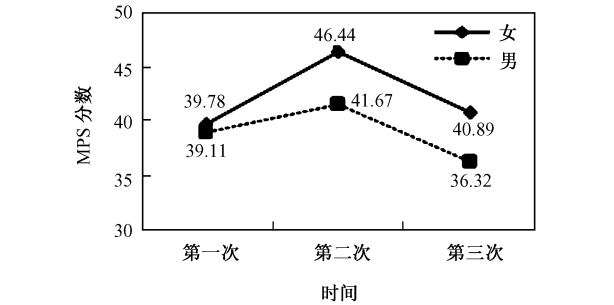


图 2 MPS×性别的交互作用分析图

### 4.5 表现性评定对不同学业水平学生 MPS 发展变化的影响

进一步考查学业水平对 MPS 发展水平的影响。将学生们 2003 年寒假数学期末统一考试成绩作为其学业水平的代表。路径分析结果表明,将实验、学业水平、实验×学业水平作为自变量(其中实验变量与学业变量均做了数据中心化处理),将对 MPS 的三次测量即 MPS1, MPS2, MPS3 作为因变量,探索其间的影响关系。由表 5 的拟合指数可以看出,  $\chi^2 = 2.627, p > 0.05$ , NFI、TLI、与 CFI 均大于 0.90, RMSEA 小于 0.05, 拟合指数达到拟合优度模型要求的标准,说明此模型与数据拟合很好。

由表 5 所示,学业水平对于三次的 MPS 都产生了显著的影响 ( $t = 16.50, p < 0.001; t = 11.52, P < 0.001; t = 11.18, p < 0.001$ ),即学业成绩越好,其 MPS 也越高。由标准化回归系数结果可以看出,随着实验的进程,学业水平对 MPS 的影响在三次测量中不断减少 ( $0.69 > 0.54 > 0.53$ )。在前两次的 MPS

上,都不存在实验×学业的交互作用 ( $t = -0.33, p > 0.05; t = 0.43, p > 0.05$ ),但在第三次时则出现了两者显著的交互作用 ( $t = -2.58, p < 0.05$ )。深入分析表明,在第三次测量时,实验班与控制班的差别受到学业水平的显著影响。学业水平越高,其受实验的影响越大,实验效果越好。由标准化回归系数可以看到,随着实验的进程,实验处理×学业水平对于 MPS 的影响在不断地增加 ( $0.01 < 0.02 < 0.13$ )。

表 5 表现性评定对不同学业水平  
学生 MPS 影响的路径分析结果

因变量	自变量	非标准化 回归系数	标准误	标准化 回归系数	t
MPS1	实验	-1.07	1.29	-0.03	-0.83
	学业	0.56	0.03	0.69	16.50***
	实验×学业	-0.02	0.07	-0.01	-0.33
MPS2	实验	-6.88	1.50	-0.21	-4.58***
	学业	0.46	0.04	0.54	11.52***
	实验×学业	-0.03	0.08	-0.02	-0.43
MPS3	实验	-6.79	1.71	-0.19	-3.96***
	学业	0.51	0.05	0.53	11.18***
	实验×学业	-0.16	0.06	-0.13	-2.58*
$\chi^2 = 2.627$		df=5	$p = 0.76$		$\chi^2 / df = 0.53$
NFI=0.99		TLI=1.00	CFI=1.00		RMSEA=0.00

## 5 讨论

### 5.1 表现性评定对于学生数学问题解决能力发展变化的影响

虽然有教育学者明确指出表现性评定对以问题解决为代表的高级思维能力具有促进作用,但是问题解决领域研究者则倾向认为无论采用什么方式,问题解决能力很难提高,它是一种需要很长时间培养、发展缓慢的能力<sup>[9]</sup>。本研究结果表明,表现性评定任务对学生的数学问题解决能力的发展具有显著的促进作用,且随着实验的实施进程,这种促进作用在不断地提高。这与 Fuch 等以小学生为被试的实验研究结果具有一致性。对口语报告资料的分析表明,表现性评定对数学问题解决能力的促进作用主要体现在理解问题、元认知与问题解决策略领域。

Bottge 曾在研究中指出,对提高问题解决能力最好的建议就是教师能够提供一种能将基本技能教授与学生问题解决的经验联系起来的教学方法,通过将有趣的问题解决经验与基本技能相联系,学生们可能会从中受益许多<sup>[10]</sup>。在本研究中,表现性评定任务本身往往就是一个有待于解决的问题。首先,所有的表现性评定任务都具有很强的现实性与新颖性,其问题都来自于学生的现实生活,与普通的数学作业有着极大的不同。这使得它有可能促使学

生更主动、更广泛、更深入地激活自己的原有经验,理解分析当前的问题情境,通过积极的分析、推论活动而对于所学习的知识与技能产生新理解。而理解的合理性和有效性又在完成任务的过程中自然地得到检验,其结果可能是对原有知识经验的丰富与充实,也可能是对原有知识经验的调整与重构。已有研究表明,当学生们解决具有新异性的问题时,可能会显示出更高的问题解决能力。这是因为较强的新异性需要识别许多生活相对少见的模式,从而可能会增加元认知对模式间关系的意识,对不同问题解决图示的激活,或包括不同的问题解决算子<sup>[11]</sup>。这些都为促进数学问题解决能力的发展提供客观上的可能性。

其次,表现性评定任务往往是一个比较“大”的任务,主要体现在背景信息丰富,需要花费较长的时间与较多解决步骤来完成。通常来说,在完成这种任务的时候,首先需要形成问题表征,即反复阅读题目,理解术语,做图表,明确已知与未知,来帮助理解问题与组织数字信息,理解问题中条件之间的关系。这也是一个从任务背景中抽取有效信息,形成问题框架的过程;第二,要围绕着问题情景制订解决问题计划,将问题分解为一系列步骤。这是数学问题解决中最关键的部分,也是元认知能力的重要组成部分;最后,才是采用各种方法尝试问题解决。最初接触此类问题、面临新问题时,缺乏相应的知识和专门的解决策略,学生们更多地是采用试误式的策略。但是,随着对表现性评定任务熟悉性不断增加、探索和学习过程的不断深入,启发式逐渐增强,选择有效问题解决策略的可能性在增加,甚至大量使用非常有技巧性的策略解决问题。这一点与我国学者辛自强与俞国良对问题的研究相吻合<sup>[12]</sup>。可以看出,这些为促进数学问题解决能力的发展提供了条件。

另外,表现性评定任务实施包括小组合作讨论与作品自评。针对作品的小组合作讨论与自评促进了学生对自己作品的反思,特别是维度评价表中的各个项目,提高了对问题解决特点的认识和对作品完成过程的监控能力。从而促进问题解决过程中的元认知能力的发展。

5.2 表现性评定对不同学校类型、不同性别、不同学业水平学生的数学问题解决能力影响

本研究表明表现性评定对数学问题解决能力发展的促进作用在不同类型学校间不存在显著差异,这意味着表现性评定任务在不同类型的学校中都取

得较好的效果,其实施不受学校类型,特别是学校教学水平、环境条件的限制。

表现性评定也同等地促进了男女生在数学问题解决能力方面的发展。与数学问题解决有关的性别差异问题经常受到关注,许多研究者认为男孩比女孩更容易在此方面获得高分<sup>[13]</sup>,女孩倾向于运用更多具体的策略,而男孩倾向于运用更多抽象的策略。Gallagher 调查了在数学问题解决表现方面的差异,发现男孩比女孩更善于运用各种策略,特别指出在问题解决方面的性别差异直到青春期才出现,其主要体现在对包含多步骤的问题解决上<sup>[14]</sup>。值得注意的是,本研究以初中生为被试,并未发现性别之间在实验效果上差异,这与以前的结果不尽相同,还有待进一步研究证实。

此外,本研究还表明,表现性评定对学生数学问题解决能力发展的促进作用受到学生原有学业水平的影响,学业水平越高,促进作用越大。这个结论与已有的研究存在着一致性,Fuchs 等的研究表明,表现性评定对不同学业水平小学生的数学问题解决能力的发展具有不同方面的促进作用,Woodward 等的研究则指出学业水平较低的学生需要更长的时期与更多的有效作业才能获得数学问题解决能力的发展<sup>[15]</sup>。结合本研究的实际情况,出现这种情况的主要原因可能在于,学业水平较低的学生由于原有的基础知识不够扎实,使得其问题解决能力的提高与发展速度受到一定的影响;其次,就是表现性评定任务为学生们提供了充分的时间与空间,对自己要求严格、自觉性强的孩子,可能会进步得更快,甚至出现两极分化的现象。因此,在实施过程如何顾及原有学业水平较差学生的发展还有待深入研究。

6 结论

(1)表现性评定对学生的数学问题解决能力的发展具有显著的促进作用,且随着实验的实施进程,这种促进作用不断增加。对口语报告资料的分析表明,表现性评定对数学问题解决能力的促进作用主要体现在理解问题、元认知与问题解决策略方面。

(2)表现性评定对学生数学问题解决能力的发展的显著促进作用不受学校类型的影响,不受学生性别的影响,但受到学生原有学业水平的影响。原有学业水平越高,促进作业越大。

参考文献:

[1]Stiggins R J. Design and development of performance assessment. <http://www.cnki.net>

Educational Measurement: Issues and Practice, 1986, Vol(6), 32—42.

[2]Haertel E H. Performance Assessment And Education Reform. Phi Delta Kappan, 1999, Vol(11), 517—523.

[3]Brenda S. A theory-based framework for assessing domain-specific problem-solving ability. Educational Measurement: Issues and Practice, 1995, Vol(14), 29—35.

[4] Brookhart M B, Durkin D T. Classroom assessment, student motivation, and achievement in high school social studies classes. Applied Measurement in Education, 2003, Vol(16), 27—54.

[5]Fuchs L S, Fuchs D, Karns K, et al. Mathematics performance assessment in the classroom: effects on teacher planning and student problem solving. American Educational Research Journal, 1999, Vol(36), 609—646.

[6]National Council of Teachers of Mathematics (NCTM). Curriculum and Evaluation Standards for School Mathematics. Reston, Virginia: NCTM, 1989, 70—71.

[7]Glaser R. Expert knowledge and processes of thinking. In D. F. Halpern(Ed.), Enhancing thinking skills in the sciences and mathematic, Hillsdale, NJ: Erlbaum, 1992, 63—75.

[8] Taylor K L, Dionne F. Accessing problem-solving strategy knowledge: the complementary use of concurrent verbal protocols and

retrospective debriefing. Journal of Educational Psychology, 2000, Vol(92), 413—425.

[9]Lester F K. Musing about mathematical problem-solving research: 1970—1994. Journal for Research in Mathematics Education, 1994, Vol(25), 660—675.

[10]Bottge B A. Reconceptualizing mathematics problem solving for low-achievement students. Remedial & Special education, 2001, Vol(22), 102—122.

[11] Brown D E. Using examples and analogies to remediate misconceptions in physical: Factors influencing conceptual change. Journal of Research in Science Teaching, 1992, Vol(29), 17—34.

[12]辛自强, 俞国良. 问题解决中策略的变化: 一项微观发生研究. 心理学报, 2003, 35(6), 786—795.

[13] Fennema E. New perspectives on gender differences in Mathematics. Educational Researcher, 1998, June-July, 19—21.

[14] Gallagher A M. Gender differences in advanced mathematical problem solving. Journal of Experimental Psychology, 2000, Vol(75), 165—190.

[15]Woodward J, Monroe K, Baxter J. Enhancing student achievement on performance assessments in mathematics. Learning Disability Quarterly, 2001, Vol(24), 33—46.

## The Experimental Study on Effect of Curriculum-embedded Performance Assessment on Mathematical Problem Solving

ZHANG Yong-mei<sup>1</sup> LIU Hong-yun<sup>2</sup> MENG Qing-mao<sup>2</sup>

(1. Beijing Academy of Educational Sciences, Beijing 100031;

2. School of Psychology, Beijing Normal University, Beijing 100875)

**Abstract:** Curriculum-Embedded Performance Assessment (CEPA) is an important approach to assessing students' learning achievement, which is highly advocated by the New Curriculum Reform in K-12 education area. Selecting seventh-grade students as its research subject, this study took a one-year longitudinal experiment, explored the effect of CEPA to students problem solving in the Mathematics. The result indicates that this assessment approach promotes students' problem solving ability significantly, and the promoting effect accelerated as the experiment progressing, which is also influenced by students' original achievement abilities. The further analysis to the subjects' oral reporting materials shows that CEPA promoting effect demonstrates in the aspects such as problems understanding, meta-cognition and problem solving strategies, etc.

**Key words:** Performance Assessment; Curriculum-Embedded Performance Assessments; Math Problem Solving