

任子朝，等：新高考数学多选题考查功能研究

发布于：强基计划网 发布人：百学 人气：1694 发布时间：2019-09-06

基金项目：国家教育考试科研规划2017年度课题——新高考不分文理科后的数学命题研究(GJK2017005)。

作者简介：任子朝(1961—)，男，研究员，主要从事数学教育及考试评价研究。

摘要：新高考实行不分文理科的科目改革，对统考科目提出了新的功能定位和区分选拔要求。数学考试必须研究、创新试卷结构和试题形式，以增强数学考试的选拔功能，实现考试目标。高考数学研制了多选题，并对多选题进行了测试。定性、定量的分析结果表明，多选题有利于区分学生，有利于提高全卷得分率，多选题的选项总数和正确选项数量将会影响学生的作答时间和得分率。

关键词：高考改革;新高考;数学考试;多选题

问题的提出

1. 高考内容改革

《中共中央关于全面深化改革若干重大问题的决定》指出高考改革的方向：探索全国统考减少科目、不分文理科、外语等科目社会化考试一年多考。在新高考中，对语文、数学、外语三个统考科目提出了新的功能定位和更高的区分选拔要求。现行高考统考科目中只有数学文理分科，不分文理科的高考改革要求对数学科提出了艰巨的任务和严峻的挑战。因此，数学新高考必须研究、创新试卷结构和试题形式，以增强数学考试的区分和选拔功能，实现考试目标。

2. 高考数学题型沿革

高考一直在进行题型改革的探索，在本世纪初，把填空题作为题型改革的试验田，进行了多选题的试验。当时的多选题给出5个备选项，但以填空题的形式出现，学生只有全部选对所有正确选项，而且没有选错的选项才能得分，对学生的考查要求是比较高的。数学多选项试题具有无需解题过程，考试分值小、考查容量大、解题思路广、数学思想丰富、对学生进行多层次区分的特点。因此，多选题对能力的考查更加深入，要求学生具备完整、细致、全面的思维品质。

国内学者对于多选题的题型结构、构成要素、基本功能都进行了相关的研究。笔者及团队成员在2015年进行了高考新题型的测试。测试对包括多选题在内的多种新题型进行了研究,统计数据和考后调查表明,新题型能够很好地发挥区分作用,有效区分学生。虽然各种题型的考查目的不同,但相互协调,能够比较全面地考查各种数学能力。余继光论述了多选题的评价目的、评价依据、评价细则,还分析了多选题的题型结构、构成要素、能力属性,认为多选题的构成要素较复杂,破解的难度较大,需要的能力点也较多,运算、推理、应用能力都可能会触及到,多项选择题增加了学生思维的复杂性,要求学生更加细心。王文雅、李玉长对高考数学中多选题的类型进行了分析,包括直接计算型、综合判断型、分析论证型、开放探索型和信息创新型等。倪兴隆、夏敬花分析了多选题型在数学能力测试中的作用,多选题突出了数学核心概念,强化了基础知识和基本技能的有效落实;关注了学生合情推理和演绎推理的有机结合;依托数学模型,注重了对数学思想方法的考查。李瑜、丁树良论述了多选题认知诊断测验编制原则、要求,以及步骤,对高考数学多选题的编制有一定的借鉴意义。这些研究都在开发多选题的题型、分析多选题的考查特点、充分发挥多选题的考查功能等方面不断深入,为新高考的题型创新设计提供了可贵的参考和借鉴。

二

研究设计

1. 测试目标

本次测试设计了有多个正确选项的选择题,即多选题,设计了考试时间为120分钟的试卷,包括了现行高考使用的单选题、填空题和解答题,以便与多选题进行比较,检验测试效果。因此,本次测试的目的之一就是测试多选题的考核效果和区分功能。测试后,请师生填写了调查问卷,并召开座谈会,听取了中学师生对新题型的反馈意见。从定性和定量两方面进行分析,为创新数学学科新高考的题型设计和试卷结构进行理论和实证研究。

2. 测试对象

本次测试在2018年3月到5月进行,对象选取广东省、山东省和浙江省的高三学生。如表1,广东省共有4 745名学生参加测试,包括文科学生2 240名,理科学生2 505名;山东省共有954名学生参加,包括文科学生386名,理科学生568名;浙江省已经实行不分文理科的高考改革,共有816名学生参加测试(注:为统计方便,在本次测试中,将浙江省816名学生全部算为理科学生)。

表 1: 三省问卷分析数据概况

省份	文科学生	理科学生	合计
广东省	2 240	2 505	4 745
山东省	386	568	954
浙江省	—	816	816
合计	2 626	3 889	6 515

新题型试卷测试时间为120分钟,学生完成测试后继续作答配套学生问卷,任课教师及阅卷教师完成相应的教师问卷。

为记录学生的作答时间,本次测试在广东的部分考场采取上机测试单选题、多选题的作答形式(在这部分考场进行测试的文、理科学生各有98名,下述表6、表7只针对这部分学生进行统计);填空题、解答题部分采取计算机呈现题目内容,学生纸笔作答,再利用高速摄影机拍照上传答案,记录学生作答时间,用于对学生每题的答题时间和全卷答题时间的分析。

3. 测试工具

本次测试命制了两道多选题,全卷试题按单选题、多选题、填空题和解答题的顺序排列,两道多选题分别为第9题和第10题。多选题的指导语为“在每题给出的选项中,有多项符合题目要求。全部选对的得4分,有选错的得0分,部分选对的得2分”。

数据分析

两道多选题选项数目均为5个，其中第9题有2个正确选项，第10题有3个正确选项。因为在测试中采用上机测试和高速摄影机记录答题时间，因此对两个试题分别从作答正确率、选项分布、作答时间等方面进行分析。

1. 学生得分概况

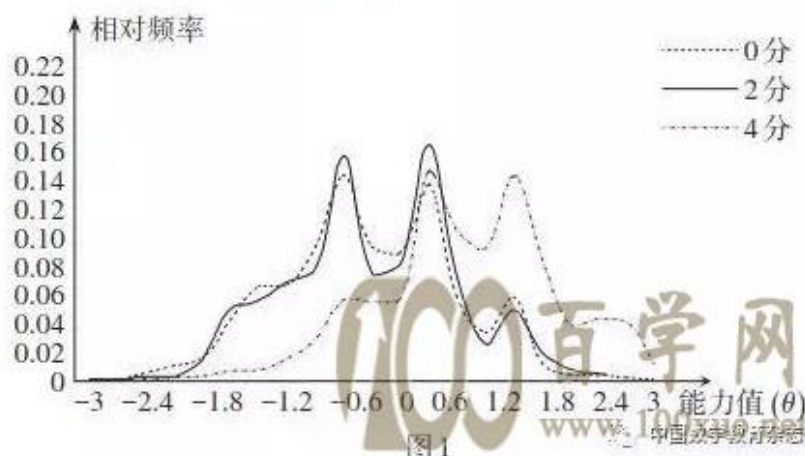
学生对第9题的作答情况如表2所示。

作答情况	全体学生	文科学生	理科学生
完全正确 (4分)	17.96%	9.38%	23.75%
部分正确 (2分)	13.06%	11.83%	13.89%
错误 (0分)	68.98%	78.79%	62.36%

从表2可以发现，大部分学生在第9题上作答错误，说明此题难度较高。对比文科学生与理科学生可以发现，理科学生在此题中得分的比例高于文科学生，且完全正确的人数比例更大。

将第9题得不同分数的学生能力作分布图(如图1)可以发现，得4分的学生集中在高能力水平；得2分与得0分的学生集中于中等能力水平，且分布相似。说明第9题能够区分出高能力水平的学生，但是对中、低能力水平的学生区分能力较弱。

第9题不同得分学生的能力分布图（全体学生）



学生对第10题的作答情况如表3所示。



表3: 多选题第10题作答情况分析

作答情况	全体学生	文科学生	理科学生
完全正确（4分）	14.47%	8.39%	18.60%
部分正确（2分）	选2个	32.74%	32.32%
	选1个	33.21%	33.44%
错误（0分）	19.58%	25.87%	25.87%

从表3可以得到，第10题得2分的学生比例较大，完全正确的学生与作答错误的学生比例较少，说明第10题难度较为合理。对比文科学生与理科学生可以发现，理科学生在此题作答完全正确的人数比例高于文科学生，作答错误的学生比例小于文科学生。

从表3可以得到，第10题得2分的学生比例较大，完全正确的学生与作答错误的学生比例较少，说明第10题难度较为合理。对比文科学生与理科学生可以发现，理科学生在此题作答完全正确的人数比例高于文科学生，作答错误的学生比例小于文科学生。

将第10题得不同分数的学生能力作分布图(如图2)可以发现，在不同的作答情况中，学生的能力分布不同，其中得4分的学生能力最高，得0分的学生能力最低，在得2分的学生中，选择2个正确选项的学生能力高于选择1个正确选项的学生。说明第10题能比较有效地区分各个能力水平的学生，但在赋分方式上有待改进，以便更精确地区分出得2分的不同能力的学生。

第10题不同得分学生的能力分布图（全体学生）

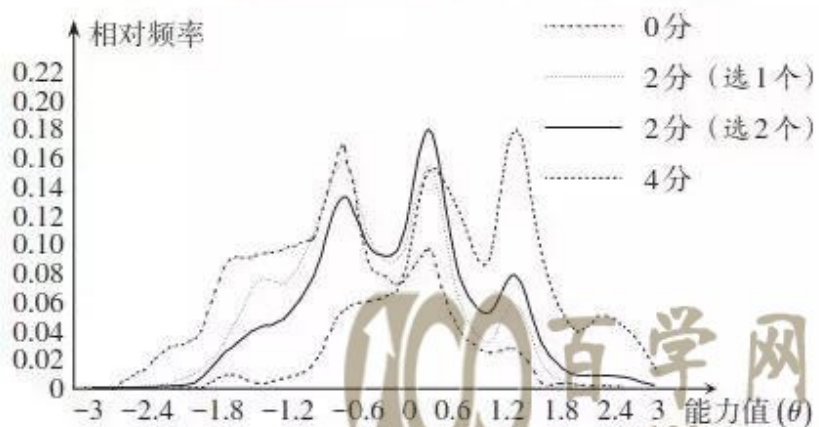


图2

2. 学生选项分布

第9题选择策略统计表如表4所示。

表 4：第9题选择策略统计表

	全体学生	文科学生	理科学生
单个选项	0.23	0.22	0.24
多个选项	0.77	0.78	0.76

学生在第9题的作答中，选择多个选项模式的学生比例较多，选择单个选项模式的学生比例较少，说明学生在此题中更愿意采取积极得分策略，争取更高的分数。

第10题选择策略统计表如表5所示。

表 5：第10题选择策略统计表

	全体学生	文科学生	理科学生
单个选项	0.41	0.43	0.39
多个选项	0.59	0.57	0.61

从表5可以发现：选择多个选项模式的学生比例较多，选择单个选项模式的学生比例较少，但两者相差较小，说明第10题中，学生更倾向于采取保守得分策略，即选择一个选项去得到2分。

3. 学生作答时间分析

第9题、第10题作答时间分析表分别如表6、表7所示。

表6：第9题作答时间分析表

第9题得分	全体学生			文科学生			理科学生		
	时间 / 秒	总分	人数	时间 / 秒	总分	人数	时间 / 秒	总分	人数
0分	140.39	54.33	143	145.24	47.09	82	133.86	64.07	61
2分	177.49	60.19	37	178.64	52.61	14	176.78	64.80	23
4分	150.58	68.34	16	109.05	46.50	2	156.51	65.03	1

表7：第10题作答时间分析表

第10题得分	全体学生			文科学生			理科学生		
	时间 / 秒	总分	人数	时间 / 秒	总分	人数	时间 / 秒	总分	人数
0分	178.78	51.52	29	127.49	43.73	15	233.74	59.86	14
2分（选1个）	205.52	54.80	91	207.79	46.82	49	202.88	64.11	42
2分（选2个）	393.71	58.10	56	374.19	50.65	27	411.89	65.03	29
4分	279.03	67.80	20	205.20	53.36	7	318.54	65.03	1

对于第9题,通过比较各得分模式的学生平均作答用时发现,在文科学生和理科学生中均为得2分的学生作答用时最多。可能这部分学生属于中等水平,力图通过多用时间,弥补自己能力的不足,争取多得一些分数。在全体学生和理科学生中,得0分的学生用时最少,可能有部分学生没有深入思考,仅凭感觉随意作答。而在文科学生中,得4分的学生用时最少,但只有两个学生,可能这两个学生在全体学生也属于水平最高的层次。平均而言,得4分的学生用时少于得2分的学生,反映出这部分学生水平较高,思维敏捷,能够很快得出答案,所以用时较少。

通过比较文、理科学生的平均作答用时发现,在得0分和得2分的学生中,文科学生的作答用时多于理科学生;而在得4分的学生中,文科学生的作答用时少于理科学生。但人数很少,说明有少部分水平很高的文科学生。

通过比较学生在第9题上的得分和整卷总分发现,理科学生在第9题上的得分与总分之间一致性较高,文科学生在第9题上的得分与总分之间一致性不够高。

此题的区分度达到0.42,综合可知,第9题的得分与总分之间的一致性较高。

对于第10题,通过比较各得分模式的学生平均作答用时,发现在文科学生和理科学生中均表现为得2分(选2个)的学生作答用时最多。原因应该与第9题相同。

通过比较文、理科学生的平均作答用时,发现在得2分(选1个)的学生中,文科学生的作答用时多于理科学生,而得0分、得2分(选2个)和得4分的学生中,文科学生的作答用时少于理科学生。

此题的区分度达到0.31,通过比较学生在第10题上的得分和整卷总分发现,无论文科学生还是理科学生,在第10题上的得分与总分之间一致性较高。

在得2分的学生中,通过比较得2分(选1个)和得2分(选2个)的学生发现,得2分(选2个)的学生作答用时和总分均比得2分(选1个)的学生多,但是现在的计分方式不能很好地区分这两部分学生。

4. 问卷调查和师生访谈

同时,通过问卷调查和师生访谈,学生和教师也都对多选题提出了意见和看法。大部分教师对引入多选题表示赞同,认为能够更精确地区分不同能力层次的学生。但是还需要在平时教学中增加训练,同时由于引入多选题,有可能会对整卷的作答时间产生影响,还希望能够控制多选题的数量。与教师相比,学生没有更多的顾虑,他们对多选题普遍持欢迎的态度,认为增加多选题后对自己的成绩不会产生太大的影响。

综上所述:(1)多选题的得分与总分之间的一致性较高;(2)在得2分的作答模式中,选对选项数越多的学生总分越高,所用的时间也越多,但是现在的计分方式不能很好地区分这些学生。

四

结论与思考

多选题的引入有利于提高全卷得分率,有利于区分学生,但选项数量设置有必要进行改进。相对于传统单选题而言,学生作答多选题时会有更多得分模式。从多选题的得分情况来看,得中间分数(2分)的学生比例较大,即多选题更容易让学生得到基础分,从而有利于全卷得分率的提升。

多选题的选项总数和正确选项数量会影响学生的作答时间和得分率。具体而言,与同等难度的单选题(4个选项)相比,多选题选项总数(5个选项)较多,学生作答时间较长;相同选项总数的多选题中,正确选项数目越多,学生的得分率越高。多选题的多级得分模式有利于提高低水平学生的得分,也有利于区分出高水平的学生。

建议在文理不分科背景下,在新高考数学中引入多选题。同时,由于作答时间保持不变,因而建议在现有多选题选项设置上减少多选题选项总数,由5个减少为4个。