

挖内力 借外力 聚合力 强攻关

——江苏省教育考试评价试点工作进入评价报告撰写攻坚阶段



10月21—26日，江苏省教育考试评价试点工作重要环节——评价报告撰写工作正式拉开帷幕，这标志着江苏省教育考试评价试点工作已进入攻坚阶段。

21日下午，我院召开评价报告撰写动员会，来自全省8个市共18名学科专家出席会议，省教研室副主任董洪亮应邀出席，北师大心理学院专家团队对评价报告数据统计进行解读。省教育考试院纪委书记吴仁林出席会议并讲话，会议由科研处处长杨从意主持。

吴仁林指出，考试评价试点工作是该院年初启动的一项全新工作，具有前瞻性、开拓性、创新性意义。开展评价试点工作既有利于中学的教育教学，也有利于地区教育行政决策，意义重大。吴书记介绍，为做好今年的评价试点工作，前期，我们深入多地开展调研，研制实施方案；成立全省教育考试评价领导小组及工作小组，强化评价工作全过程开展的统筹协调力度；综合考虑、科学分析、合理确定苏州、泰州两地共四所中学作为评价试点对象；科学研制评价指标体系；与北师大心理学院签署评价合作框架协议，立足高起点，提升评价工作专业性。目前，北师大心理学院团队已完成撰写数据描述部分评价报告。吴书记指出，如果说，前期的工作属于定量分析，接下来工作就是在定量基础上的定性。为了撰写高质量的诊断性建议，推出高质量的评价报告，我们着力在专家团队的遴选方面下功夫，经过多方推荐、综合考虑、严格把关，最终选定18位在本学科领域有较高知名度

（上接第三版）高中，是否合理还有待进一步的检验；②考试内容并没有体现选考科目70%学考模块加30%选考模块的要求，区分度不一定能体现选考科目考试的要求；③只考了物理等五门选考科目，缺少对生物和技术两门课程的测试，对于考生的科目选择以及最后的模拟结果可能会有一定的影响；④对第二次

和影响力的学科带头人或专业研究者参与撰写团队，共同完成撰写任务。吴书记强调，吸纳学科专家参与撰写评价报告，可以借助学科专家丰富的教育教学经验和对教学现状的全面了解进行评价，并对问题间的关联和因果关系进行深入剖析，从而实践教育评价的基本原理，使评价建议更有温度、热度、深度、效度，更能够被有效地接受和采纳。

吴书记希望专家团队能强化责任、团结协作，多建睿智之言，多献务实之策，凝聚团结之力，确保高质量、高要求完成撰写任务，为今年考试评价试点工作乃至以后的教育考试事业发展打下坚实的基础。

董洪亮作会议发言，董主任充分肯定考试评价试点改革这一创新举措的现实意义，强调开展评价建议撰写的重要性以及必要性，希望学科专家能将此次撰写任务与教育教学实践相结合，充分利用自身优势，高质量高标准完成好撰写任务，为今年全省教育考试评价试点改革作出贡献。

动员会结束后，学科专家随即投入工作，对2个省辖市、4所中学高考考试数据撰写诊断性建议，截止26日上午，6份报告共10万余字撰写任务已经顺利完成。11月中下旬，我院将与北师大心理学院专家团队一起赴试点地区与学校进行报告解读、交流与研讨，为完善高考评价指标体系、扩大试点范围听取建议、积累经验。



考试成绩的模拟只是根据特定假设进行简单估算，这种估计与实际结果相比不免会有一些出入。

——本文摘编自《中国高教研究》2015年第6期。作者简介：文东茅，北京大学教育学院/教育经济研究所教授；鲍旭明，浙江省慈溪中学校长、中学高级教师；傅攸，浙江正睿教育科技有限公司副总经理、博士。

江苏教育考试科研月报

2016年第10期（总第22期）

江苏省教育考试院编印

要闻传递

我院成立江苏省教育考试学术委员会

10月27日下午，江苏省教育考试学术委员会成立暨第一次全体委员会议在南京隆重召开。学术委员会主任委员、院长鞠勤出席会议并讲话，首届学术委员会全体委员参加会议，会议由学术委员会副主任委员、院纪委书记吴仁林主持。

成立会议宣读了首届学术委员会主任委员、副主任委员、委员名单，鞠勤院长向新一届学术委员会委员颁发聘书，并与全体委员合影留念。

鞠勤院长在讲话中首先对首届学术委员会的成立、对当选的各位委员表示祝贺。他指出，当前，全省教育考试科研工作取得显著成绩，全省招考

系统致力于科研工作的积极性和主动性不断增强，科研能力和整体素质不断提升。为进一步加快全省教育考试科研事业发展，规范学术研究、促进学术民主、加快学术指导，充分发挥专家在科研决策中的咨询和参谋作用，省教育考试院成立了江苏省教育考试学术委员会，学术委员会的成立，可以更好地为教育考试发展决策服务，可以提高教育考试科研发展水平，可以培养和吸引人才并增强科研团队整体素质，意义重大。当前，我省正处于新一轮高考改革方案研究关键之年，学术委员会的成立恰逢其时，既是大势所趋，又是迫切需求。希望全体委员能够依据章程规定，认

真履行职责，积极建言献策，为促进全省教育考试科研学术繁荣和整体管理水平提升贡献智慧和力量。鞠院长还表示，将在经费投入、人才队伍、职权行使、机制运行等方面提供支持，为学术委员会充分发挥作用提供保障。

全体委员随后发言，对我院学术委员会的运行机制、制度设计，以及如何更好发挥作用建言献策。委员们从六个方面提出建议，一是要追求高位目标，力争推出高质量科研成果；二是要凝练选题方向，突出问题导向、目标导向，加强实践性研究；三是要拓展职责范围，不仅服务科研事业发展，更要为全省教育考试事业作出

贡献；不仅要加强学术研究，还要加强人才队伍建设；不仅研究微观课题，还要加强宏观研究，服务决策咨询；四是要突出重大重点课题研究，将研究与实操有机融合，以学术助推事业发展；五是要创新工作方式，丰富工作形式，推进智慧碰撞；六是要加强学术队伍建设，充分吸纳多领域人才加入研究队伍，凝聚力量合力攻关。委员们一致表示，会珍惜荣誉、履行职责、积极为江苏教育考试事业改革和发展作出自身贡献。

成立大会结束后，全体委员随即开展工作，对全省第五届科研论文进行终评，确定最终获奖名单。



学术交流

编者按：浙江高考改革的重要内容之一是选考科目实行等级赋分制，不少人担心这种计分方式将导致大量同分现象，影响高考招生的区分度。本文利用浙江省九所高中联考的数据，模拟出自选科目和等级赋分之后浙江可能出现的自选科目满分人数，并对等级赋分可能出现的问题和对策进行了讨论，供读者参考。

等级赋分对高考区分度的影响

——对浙江“九校联考”数据的模拟分析

文东茅 鲍旭明 傅攸

■ 问题的提出

浙江高考综合改革试点方案的重要内容之一是高考科目由语文、数学、外语3门必考科目和另外3门选考科目构成，其中选考科目实行等级赋分制。等级赋分制对高考成绩的分度到底有何影响？比较有说服力的方式就是利用适当的数据进行模拟分析。

■ 数据来源和模拟分析方法

（一）数据来源

本研究利用的是分布于浙江省九个地市的九所重点高中组成的“九校联盟”于2015年3月进行的一次统一测试数据。参加本次考试的是各校2014年入学的所有高一学生，考试科目包括语文、数学、外语3门高考科目和物理、化学、政治、历史、地理等5门选考课程。本次统考总人数为7350人，实际用于统计分析的考生数为7298人。

（二）模拟分析过程和方法

1.通过“5选3”为每位考生确定选考科目。对所有考生各科成绩进行排名，并对相对名次进行比较，选出相对名次靠前的三科作为该考生的选考科目。

2.为每一位考生的选考科目进行等级赋分。①计算出一本上线率为70.5%。②按浙江高考改革方案规定选考科目前15%的学生被划为5等赋分的比例。③计算出各科成绩的累计百分比，按照“最接近累计比例”原

则划定各个分值对应的累计比例和原始分值。④根据转换关系对每一位考生的每一科成绩进行等级赋分。

3.统计第一轮考试之后学生的成绩状况。假设本次考试为考生第一次参加选考科目的成绩，对各位考生三门选考科目成绩进行加总，统计各个分数段的学生人数比例；同时统计“5选3”构成的10种科目组合的人数比例和高分段学生人数。

4.通过统计概率对考生第二次考试满分状况进行模拟。估算第二次考试后选考科目获得300分的人数比例。

■ 模拟结果

（一）学科选择及学科组合情况

根据排名优先原则为每位同学“5选3”之后，发现各科目的人数比例比较接近。通过“5选3”，可形成10种科目组合，按照“排名优先”的逻辑，各个科目组合都有一定比例学生，但人数比例有一定差别，传统上偏文科的政历地组合的人数比例最高15.9%，物政地的组合人数最少。

（二）等级赋分情况

在进行等级划分的模拟时，在几乎每一科目的每个等级上都遇到了“归整困难”，从表1可以看出，几乎所有等级都需要按照“最接近”的原则取值，这是等级赋分需要面对的常见问题。

表1 各科等级赋分比例及对应的原始分

预定值		物理		化学		政治		历史		地理	
赋分	累计 %	考分	累计 %	考分	累计 %	考分	累计 %	考分	累计 %	考分	累计 %
100	4. 67	90	4. 2	97	5. 5	91	4. 6	92	5. 1	94	5. 8
97	14	83	14. 5	95	14. 0	88	13. 6	89	15. 0	92	13. 2
94	28	78	26. 5	93	25. 6	85	29. 2	86	30. 5	89	30. 1
91	46. 67	72	46. 2	90	45. 0	82	45. 6	83	49. 4	86	49. 0
88	70	65	69. 1	85	70. 2	77	69. 6	79	71. 2	82	71. 6
85 及以下	100	14	100	9	100	21	100	10	100	26	100

（三）第一轮考试之后各科成绩统计

从中可以看出，各个学科100分的人数差别较大，从181人到256人不等。各科总计有1106个100分，1982个97分，分别占考试总门次的5.05%和9.05%，而不是预定的4.67%和9.33%，这种差别也是由于按“最接近累计百分比”划线的误差造成的。

表2 第一次考试各科各分值人数统计

赋分	物理	化学	政治	历史	地理	合计
100	181	239	204	226	256	1106
97	442	371	395	442	332	1982
94	519	505	686	685	744	3139
91	848	845	722	841	828	4084
88	985	1096	1056	968	991	5096
85 及以下	1332	1229	1335	1279	1247	6422

为考生的选考科目赋分之后，通过将三门科目成绩加总可以获得考生第一轮考试之后的总成绩。在此次模拟中，三门选考科目总分最高为300分（即三科均为100分），最低为255分（即三科均为85分），在此之间每3分为一个分值，由此分出16个分值（在浙江省的实际考试中，将会在120至300分之间分出61个分值）。

表3列出了其中最高的4个分值的人数分布。从中可以看出，如果以此次统考作为第一轮学考成绩，仅有38人三科均为满分，占九所高中考生人数的0.5%。以此估算，在全省学生中第一次考试三科满分的只有约0.1%，按30万考生计算，仅有约300人。

表3 不同科目组合人数及高分数人数分布

科合	人数	300 分	297 分	294 分	291 分
物化政	907	2	7	17	19
物化史	686	5	10	13	11
物化地	927	9	5	13	21
物政史	574	2	11	7	13
物政地	526	2	5	7	16
物史地	681	6	7	12	19
化政史	735	5	8	10	22
化政地	494	2	5	9	13
化历地	603	2	9	21	8
政史地	1157	3	8	16	31
合 计	7293	38	75	125	173

（四）对第二轮考试高分状况模拟

通过第二轮考试，两轮考试总分变为300分者为101人，占考生总数（7298人）的1.38%。以全省30万

学术交流

考生计算，大约有889人可以获得三科满分。由此可见，在等级赋分之后，三科满分的人数比凭简单估算得出的“数千人”要少很多。

■ 对等级赋分相关问题的讨论

（一）关于同分人数过多问题

问题：通过本研究的模拟，浙江选考科目有大约900人可获得三科满分；尽管选考科目满分者比估计的少，通过语数外成绩也有足够的区分度，但这并不能完全打消人们的质疑，因为在此情况下各个专业招收的不是相关科目上有优势、有潜力的学生，而仅仅是语数外基础较好者。

建议：可以在高考考完语数外之后进行一场加试，允许考生在选考科目获得满分（或放宽到97分甚至94分）的科目中选择最多一门科目参加考试。该科成绩不计入总分，而是作为高校对考生进行综合素质评价的参考，或者部分替代高校自主招生的笔试，如果没有增加加试，笔者也建议同时向高校提供等级分和原始分，以增加分数的区分度和招生的依据。

（二）关于在等级赋分情况下是否要参加第二次考试问题

问题：本研究的模拟，是假定第一次考试满分者将不参加第二次考试，而其他所有没有获得满分者都将参加第二次考试。在实地调研访谈过程中，被访问的师生也都表示，只要第一次考试没有得到满分，第二次考试一定会选择参加，甚至已经获得满分者也有可能为了“占位”也选择参加第二次考试。

建议：引导考生理性认识绝对分（原始分）的重要性，认识到在第一次考试中已经发挥了水平的考生放弃第二次考试，一定程度上减少选考科目满分者的比例。对于第一次考试已经是满分者，可以制定政策，禁止他们再参加第二次考试，以引导良好的竞争文化。

（三）关于等级赋分与原始分对不同考生的影响

问题：长期以来，学生习惯于参考原始分来评判自己的学习状况和相对位次，等级赋分则可能出现以下情况：①将会给予不同原始分相同的分值。②将扩大或缩小同一科目原始分之间的差距。③对偏科程度不同的考生影响不同。

建议：应该及早引导考生正确认识这种赋分方式不同带来的变化；为避免对考生产生不必要的心理影响，建议不向考生公布选考科目的原始分。

■ 局限性

①所利用的数据来源于九所特定（下接第四版）