小学考试的那些事

——基于核心素养的考试与评价改革

专业学位类别：教育

专业领域：基础教育

适用课程：教育测量与评价

作者姓名：于海英、甄莹

工作单位：牡丹江师范学院

小学考试的那些事——基于核心素养的考试与评价改革*

摘要：学生发展核心素养是落实立德树人根本任务的一项重要举措。小学考试作为义务教育基础学段教育教学质量的测量与评价手段，在确保学生核心素养得以发展方面有着重要的导向功能。B 小学经历了出题权力由全市统考、区里统考，最后下放到学校的发展历程。期间除了出题人发生了变化，考试试卷编制依据、试卷的评阅程序和环节、考试质量分析用途并没有太大变化。B 小学及 C 校长在考试演变发展中的经历与困惑，促使人们思考小学考试的设计与实施应该考虑核心素养的教育要求与社会期望，以及将核心素养融合到考试中的有效路径与策略。

关键词：核心素养、小学考试、效度、教育目标

The Effectiveness of Primary School Examination under Core Literacy

Abstract: The core literacy of student development is an important measure to implement the fundamental tasks of morality education. Elementary school examinations, as the means of measuring and evaluating the quality of basic education in compulsory education, have important guiding functions in ensuring the development of students' core literacy. The B elementary school went through the development process of the city's unified examination, the district unified examination, and finally the devolution to the school. Except for the change of the person who made the question during the period, there was not many changes in the basis of the content of the test paper, the review procedures and links of the test paper, and the use of the quality analysis of the test. The experience and perplexity of the B elementary school and principal C in the evolution of the examinations have prompted people to think about the design and implementation of the primary examinations. The education requirements and social expectations of core literacy should be considered, as well as effective paths and strategies for integrating core literacy into the examination.

Key words: core literacy; elementary school exam; validity; educational objectives

---

*作者简介：1. 于海英，女，黑龙江省牡丹江市人，牡丹江师范学院教育科学学院教授，主要从事教育测量与评价研究；2. 甄莹，女，黑龙江省佳木斯市人，牡丹江师范学院教育科学学院助教，主要从事教育测量与评价研究。

编制说明：按照研究的要求，也是处于对本案例中涉及人员的尊重，作者对案例中涉及学校的名称、所在城市及相关人员等，做了必要的掩饰性处理。

背景信息

在当今社会的境遇之下，深入思考“教育要培养什么样的人”这一根本问题时，相信各国的答案都会是一致的。以往以知识和技能为中心、以“量”的扩充为追求的教育模式不能与复杂多元且急剧变化的时代需求相契合，只有找到学生终生受益的点，注重教育“质”的提升，培养具有终身学习能力、合作能力、创新能力、国际视野等关键素质的复合型人才，才能使个体在成功融入当前及未来社会的同时促进国家的和谐与强盛。在这样的时局下，核心素养应运而生，并在联合国教科文组织、经济合作与发展组织等国际组织的影响下，在 PISA 等国际性教育测评项目的推动下持续发酵。至今，已有十几个国家、地区和跨国组织陆续建构起了本土性的核心素养模型并以此指导教育实践，核心素养俨然成为世界教育改革的支点与中心。

学生发展核心素养概念在我国的提出与应用整体上转换了教育领域的话语体系，划分的文化基础、自主发展和社会参与三个方面及其具体细化的 18 个基本要点，其根本出发点是强调社会责任感、创新精神和实践能力，促进学生全面发展。它要求聚焦课堂实践，聚焦对学生核心素养的培养。考试与评价改革是核心素养落地的重要抓手，基于核心素养的考试与评价改革能够充分发挥素质教育的引导作用。然而，核心素养在我国的正式落地距今已有几年的时间，课程目标与教学方式的转变随着课堂教学的改革逐步做到了转型式发展，而作为教育制度中“牵一发而动全身”的关键一环的考试评价，则以其高利害性的特点，争议颇多。尤其对于作为基础教育起点的小学教育阶段，学生的个性、能力、品格发展等都由此起步，即决定了学生核心素养应由小学阶段为起始。面对小学没有升学竞争任务的现实，不禁引发我们的思考——核心素养会促进小学考试的改革吗？

本案例选自 2019 年 10 月，M 市 A 区 B 小学的 C 校长对小学考试问题的访谈与调研。C 校长，B 小学教学校长，中学高级职称、市级教学能手，省级语文骨干教师。毕业于师范类学校师范专业。1991 年毕业后在一线从事小学教育工作，先后被提升为年段主任和教学校长。B 小学教育教学质量在 A 区属于中等偏下水平，学校有学生 406 人，教职工 40 人。学校教学设备完善。但学生群体中留守儿童和外来务工人员子女所占比例较大，达到 73%，与学业成绩相比，学校领导和老师对学生的心理和生活关注的比较多。C 校长是一个做事非常踏实的校长，自从小学考试权力下放之后，学校期末考试卷的编制、审核、批阅、考试质量分析等一系列关于考试的工作都由她来负责和组织。虽然在访谈与调研过程中，她并没有说出什么高深的理论，但从她的言谈举止间看到了她对核心素养促进小学考试改革趋势的认可和实践。

案例正文

1.小学考试试卷编制权力逐渐下放

2012 年度的全市教育会议的现场。C 校长和其他学校的校长们正襟危坐地在台下认真地听着教育局领导的年度工作总结：“2012 年度小学考试成绩排在前 5 位的学校是：H 小学、J 小学、K 小学、L 小学、M 小学；考试成绩排在后 5 位的学校是：A 小学、B 小学、C 小学、D 小学、E 小学……”，B 小学的 C 校长这次又因学校考试成绩排在最后 5 位被点名了，看到排在前 5 位被点名的学校校长们嘴角上扬，难掩喜气的表情，C 校长心中略微有点恨意和失落，后面的会议内容她几乎不知道领导说了些什么。会议结束后 C 校长走出会场，心里正在想着被点名的事情，听见后面有人喊她的名字，原来是和她一样被点名的 A 小学的 P 校长。A 小学的 P 校长说：“总在全市教育大会上因为成绩倒数被点名的滋味真不好受啊！”，C 校长看看 P 校长感叹：“是啊，我们学校这都连续 3 年因为成绩倒数被点名了”；A 小学的 P 校长接着说：“全市统考，所有小学都用同一标准来衡量教学质量，因为是在同一时间、使用同一份试卷，使用同一参考答案，倒是解决了一些暗箱操作的问题，但那只是在形式上实现了教育公平。怎么判定学校的好坏呀？不还是靠成绩？像我们这样的农村学校，怎么和市区学校比啊，愁死人了！”。刚从会场出来的 S 小学 W 校长听到两人的对话，马上回应道：“现在啊，怎么看你学校是不是好学校？全市学校成绩一出，拉个单儿，横向纵向一比，这顺序市教育局就给定好了”，W 校长忽然停下话，向西周看了看，很神秘地压低声音说：“有的学校为了排名靠前啊，可是花了大力气了，我听说啊，有的学校到了期末，学生啥活动都没啦，就刷题，还有和市教研员拉帮结派的。我告诉你们，现在老老实实办学校搞教育，不行！”。C 校长听着两位校长的话，觉得有点道理，心里盘算着在不违背职业道德和良心的情况下，怎么做明年才能不再被因为成绩倒数而被点名。

2013 年度的全市教育会议上，C 校长还是如往年一样因成绩倒数被点名。教育局领导在工作报告中提到了一个新的决定：“在广泛征求学校、家长的意见，经市教育局领导集体研究决定，市教育局决定取消全市统考，今后小学段考试出题的权力下放都到各辖区教育局”。虽然被点名，但是听到取消全市统考这一消息，C 校长内心产生一丝欣喜。

2014 年是全市推行教育考试改革的一年，从这一年开始，考试出题的事情都有全市各区教育局负责组织。6 月初，C 校长接到所在区教育局的通知，让他们在期末考试前一天来教育局领取所有年级的各科考试卷。6 月 24 日，C 校长带着两个段主任到教育局把期末考试卷取回，期待着明天的考试一切顺利。6 月 25

日早上 6 点半左右，C 校长就来到学校，刚进校门就听到 6 年级 3 班的 R 老师和她打招呼：“C 校长，这么早就来了”，C 校长转过头：“今年区里第一次出题，教育局还不让提前看卷子，也不知道怎么样，心里没底，今天早点过来，在考试题发给学生之前，我先看看有没有问题”，R 老师说“区里出题肯定不会有问题。”，C 校长笑着说：“理论上不会有问题，走，你和我一起到大队部”。7 点，所有的班主任都到齐了。“学生考试让咱们来这么早干嘛？”，“往年都是 7 点半来，今年为啥这么早？”，几位班主任私下里聊着天，语言中带着点抱怨的成分。“各位老师，今天这么早让你们来就是让你们在学生正式考试之前，每一位老师把自己班级的语文和数学卷子做一遍，在 7 点 40 分之前完成，如果有问题记下来一起交给我”。C 校长说完，将卷子按照年级分给了不同的班主任。7 点 40 分，班主任们将自己做的卷子存在的问题交给校长：“C 校长，我们年级的语文卷子，有多处错别字”，“C 校长，我们年级的数学卷子，题都串格了”，“C 校长，我们年级的卷子印错了”。C 校长心里产生了疑惑：“怎么会有这么多问题？”，她马上给区教育局打电话：“卷子怎么这么多问题？8 点半要考试了，怎么办呀？”，接电话的是区教育局年龄较大的语文教研员：“我们也没有办法，全区共有 9 所小学，3-6 年级语文、数学都由区里统考，一所学校共计要编制 16 套试卷，而且区教育局还让编制城市学校试卷和农村学校试卷两类，共需要编制 32 套试卷。我们教研员只有 2 人，工作量那么大，出点错误在所难免的。你把卷子里出现的错误，在电话里告诉我，我看哪有问题，告诉你们怎么改”。在 8 点 30 分考试之前，试卷中的所有错误都有了修改的建议。考试结束后，老师们把考试卷统一送到了大队部，“今年考试题出的太差了”，“我们班学生反复问错的地方”，“这种卷子感觉太随意了”，老师们边整理卷子，边讨论着。这种状态一直持续 2019 年。

2019 年初，C 校长迎来了她所认为的 M 市小学考试的春天，接到上级通知，由于区统考的反馈意见较大，同时 A 区教育局也考虑到教研员工作量太大，而且教研员容易忙中出乱的现实情况，在征求所有小学意见的基础上，将考试试题编制的权力再次下放，由各个小学自己承担每学期期末考试试卷的编制工作。从 2019 年 6 月开始，A 区教育局所辖所有小学每学期的期末考试都由每所学校自己负责编制试卷，并组织考试。C 校长由此开始了学校命题考试的新时代。

2019 年 5 月，A 区所辖 9 所小学，在位于市区的 U 小学里自发地组织了第一次出题合作会，由市区的 U 小学校长主持：“各位同仁，今天的会议的主题就一个：期末考试合作出题的事。我们大家都知道，咱们每个学校的考试科目太多，难以实现全部科目都由教学校长一人完成，所以我提议，咱们 9 家合作出题，这样工作量会少很多，大家谈谈各自的想法”。 C 校长首先发言表示：“学校间合作的形式，校间试题互用，9 所学校的期末试题都有保证，并且分摊到各位校

长身上的工作量也不大，还有时间进行检查，能够确保试题不出错误，减少后续的麻烦”，“我同意合作出题”“我同意C校长的意见，合作出题”，9位校长中的8为校长纷纷表示愿意合作出题，唯独F校校长没有发言。大家都把目光移到F校校长身上，想听听他的意见是什么：“我想参加合作出题，但是我没有办法派人出题，我们学校没有教学校长。”，“这简单，你们就不用派人出题了，我们8家出完题你用就行了呗，大家看这样行不行？”，作为主持人的U小学校长先发话了，“没问题”，“可以”，“你们学校就别出题了”，其他校长都对U小学校长的建议表示赞同。2019年6月末的期末考试试题由8所学校的教学校长负责出题，没有再听说考试卷发生错误的事情。

2.小学考试试卷编制内容的选择

2019年春季学期的期末，B小学的C校长作为出题人负责三年级语文的出题任务。因为是第一次出9所学校共用的三年级语文试卷，并且是权利下放后的第一次出题，让C校长倍感压力，在出题那段时间，C校长的办公桌每天都被堆得满满的，教材、教师用书、课程标准、往年试题。

2019年6月10日晚九点，学校早已褪去了一整天的喧嚣与热闹，陷入宁静，然而远远的就可以看见B小学C校长的办公室的灯亮着。“出题多简单的事情，你怎么出个题这么费劲呢，这都好几天了还没出完呢”，C校长的老公已经陪她熬了几个晚上了，今天实在有些受不了了，也发起牢骚来。C校长从出题的材料中抬起头对着老公缓缓地说：“出题可不简单，是一件非常难的事情，要保证出题的质量，就得做很多工作，比如，要先明确具体的学科内容，像我出三年级语文题，就得把三年级语文的教师用书和教材先看明白，然后再参考区里往年编制的试题的权重选择内容和题型。”，“什么内容、权重，我觉得你就是太认真了，把书上学过的东西出成题就行呗”，C校长打趣道：“出题啊，这里面的门道大了去了，目标、内容、权重是必要考虑的。我以前一直就没弄明白这个考试内容的权重怎么定位，就是权重这个词儿，我也是2000年参加市里教学评价培训时候才第一次听说，才知道是怎么回事，慢慢摸索，以前出题啊，是这找一道，那找一道，摸着石头过河！就是现在我也还是没有搞得十分清楚，这不，也还得照着往年市里出的题看！”，C校长的老公走到她的办公桌前，拿起C校长办公桌上的一沓资料中，最角落的《课程标准》问道：“你总说课标课标的，是这个吗？这个出题有用吗？”，C校长此时端起了她十足的教师范儿，一脸认真的给她的老公上起课来，“所谓课标，即课程标准，是教材编写、教学、评估和考试命题的依据，是国家管理和评价课程的基础；它体现了国家对不同阶段的学生在知识与技能、过程与方法、情感态度与价值观等方面的基本要求，规定各门课程的性质、目标、内容框架，提出教学和评价建议。不过出题时，这个课标就是一个参

考，主要还是教师用书和教材作为出题依据。我觉得小学三年级语文里的基础知识最重要，所以分数占 50%，再一个写作部分也比较重要，分数占 35%，其次就是段篇章，分数占 15%。”，长篇大论的一段演讲结束，C 校长自己忍不住笑起来，坦言说：“我自己的学术水平不高，对课程标准也不咋理解，也没时间研究，我知道课程标准是指导性的，但要想依据它来出题，难度太大了，理解起来太费劲，估计学生答起来也费劲。所以出题主要还是固定用教师用书和教材上的东西，同时借助其他参考资料的题型”。

2019 年 6 月 16 日上午，D 校长办公室。C 校长作为出题人与 D 校长核对自己所出的试题。D 校长看完试题后，提出一个建议：“C 校长，你代表咱们学校出题，一定要确保质量，不能出错误。另外语文出题一定要侧重语文知识和人文素养”。“D 校长，你说的是核心素养吗？”，C 校长为了确认反问了一句，D 校长说：“就算是吧，语文素养是一种以语文能力为核心的综合素养，其要素包括语文知识、语言积累、语文能力、语文学习方法和习惯，以及思维能力、人文素养等。《课程标准》总目标中所列的 10 条，是对接受义务教育的每一个学生在语文素养方面的基本要求。如果达到了这 10 条要求，就应该认为全面提高了学生的语文素养。我们可以在试题中将课程标准中的语文素养方面内容通过不同题型体现出来”，听了 D 校长的话，C 校长赶紧解释：“我出题时候还真没多想，但是我觉得我做的比较好的地方是，编题时注重将题的内容与生活联系起来，就是想侧重于引导学生解决生活中的现实问题，与实际生活结合进行出题。像口语交际这类试题的，主要就是设置生活中的日常交往活动情景；再比如，读句子一般是融入到听说读写片段中，这样就会让学生感觉考试不是陌生的事情，而是与生活更接近的事情。这应该就算把核心素养加进去了吧？”，D 校长笑着说：“应该算”。“其实我自己挺愿意在考试中试着实现核心素养融合的，但在咱们这样的学校好像不太适合，主要是太麻烦，咱们自己也拿不出一个检测用的标准，而且有的话可能也用着不顺手。我们出题与往年也不能差别太大，要考虑教师的教学实际和学生的学习状况，更重要的是还需要给家长一个满意的交代”。两位校长核对完试题后，C 校长将试题的电子版的格式、文字、标点符号、排版等各方面都进行了再次审核，无误后保存到电脑中待期末考试共 9 所学校三年级语文考试用。8 所小学教学校长联合编制的各年级各科的试卷在 2019 年 6 月末在 9 所区属小学考试的考试使用了。

3. 小学考试试卷评阅工作的开展

2019 年 6 月 22 日，B 小学组织全校教师召开了期末考试工作会议。“考试都考这么多年了，还有啥新要求吗？”，“今年考试批卷和往年一样吗？”，开会之前老师们都在议论这次会议的内容和主题。会议开始 D 校长进行了讲话：“我

们今年的期末考试与以往考试不同。这是第一次学校自己出题考试，大家一定要在遵守学校的要求，千万不能出错，下面请C校长和大家说说期末考试要求”。接着C校长提出了期末考试要求：“为保证评分客观、公平问题，咱们学校期末考试阅卷一定要按照要求进行，具体要求就是：第一，沿袭原来市里统考和区里统考的评阅程序。第二，所有考场的安排不能按照原始班级进行布置考场，要将不同的平行班级的学生混在一起重新安排考场，每一个考场中都有不同班级的学生。第三，考试的试卷学生答完后将卷头进行密封，所有学生的姓名、班级等信息都是保密的。第四，期末考试每个考场由两位老师监考，监考教师学年间交换，3年级老师到4年级学生考场监考，4年级老师到5年级学生考场监考，5年级老师到6年级学生考场监考，6年级老师到3年级学生考场监考。第五，学校会为每个考场准备一些必要的装订材料。在考试结束后，监考教师在考场就直接把试卷装订密封好统一送到大队部。这样做能保证老师和学生不能作弊，对大家来说都比较公平，不会被人家议论”。对于C校长的要求，老师们在会上没有表现出异议，会后大家也都觉得这样对老师和学生来说更公平。

2019年6月26日，B小学的期末考试如期而至，按照前期的布置与安排有序进行。早上7点30分，监考老师们都准时到校。彼此之间就监考问题寒暄着：“G老师，你和谁一起监考？”，“和S老师，你和谁？”，“我和Q老师”。8点10分所有监考老师都拿到试卷，C校长再次强调：“各位老师，一会到考场，在发卷子之前，先检查一下是否有错误然后再发”，老师们带着明确的要求进入了考场。考试过程一切顺利。考试结束铃声响起，考完试的孩子们轻松了，哼着小曲儿蹦蹦跳跳的离开学校，而对于C校长和B小学的老师们来说，有一项艰巨的任务正等着他们——阅卷。

在阅卷开始之前，C校长再一次把老师们召集起来开会，再一次强调：“为了保证评阅试卷工作的快速完成，安排每一学年的各班的班主任和科任教师评阅试卷，实施流水，其中班主任负责主观性试题的评阅，科任教师负责客观性试题的评阅，一般是3-4名教师负责一个年级的考试试卷的评阅。每人从头到尾一直负责全学年的所有试卷中的1-2道题，这样评阅的标准能够保持一致。教师评阅完每道题，将分数填写在试卷的相应位置。待一个学年所有老师都评阅完，大家一起再复核一遍所有试卷，没有问题后，拆开密封线，将学生考试分数按班级登录。每个班主任拿到自己班级的成绩再对学生成绩进行核对一遍，以确保分数准确无误。”，S老师提出自己对批卷的疑问：“要是批错了怎么办？”，C校长说：“各位老师，批卷要谨慎，试卷上的成绩落下之后基本就不能改变，除非有个别离谱的分数，这种情况我们要请示大校长怎么处理。最好不要出现这种情况。班主任评阅作文题一定要把控住分数，抓住评阅作文的关键踩分点。卷子批完后再加上还有两次复核，评阅教师核对一次，班主任再核对一次，这样试卷中分数

出错的情况不太多。”。Q 老师接着问：“作文题怎么给分呀？可能我们每个人理解的都不一样”。C 校长说：“我知道，最难的就是阅读题和作文题。一是因为这一类型题是主观题，标准不太好把握，二是作文的评阅花费的时间和精力较多。因此，为了保证阅读题和作文题的评阅客观以及教师对评阅工作没有意见，在阅读题和作文题评阅方面先制定评阅的程序。例如，就作文题评阅，将所有考场的试卷都拿过来，大致翻阅作文答题情况，找到初步判断为最好和最差的各 4-5 篇作文，然后和其他班主任、科任教师一起阅读，做出判断，如果大家觉得可以定为一等作文和最低等作文就确定下来。作文分数共定为 5 挡，一等分数最高，五等分数最低。在这里，我先把丑话说在前面，大家一起把标准订立下来，然后按标准进行评阅。如果订立标准时没有异议，过后不允许再提意见。如果要是还有意见的话，下一年谁有意见谁就自己评阅这类题型”。“哈哈，这样可以”，大家笑着表示同意。2019 年春季学期期末考试工作就在 C 校长的统筹协调、学校前期的一系列准备和老师们的共同努力下圆满完成了。此时的 C 校长是幸福并充满成就感的。用她自己的话说这次学校自己出题的试卷评阅没毛病。

4. 小学考试的质量分析报告撰写

期末考试批卷工作结束了。上午 9 点，在 B 小学的大队部，23 位老师围坐在中间的会议桌旁谈论着这学期学生的考试成绩，“我班学生这学期太给我长脸了，我班数学今年考得超乎我的预期，平均分 100 分，哈哈”，Z 老师笑着说起自己班级的数学考试成绩，那表情不亚于买彩票中了大奖；旁边的 F 老师情绪不高地说：“我班数学满分的学生也挺多，要不是有两个拖后腿的，也都能满分”，N 老师插了一句话：“哪两个拖后腿的？”，“哎，就是那两个呗，一个是爸爸娶了后妈，后妈只管给他做饭，其他什么也不管的那个学生，另一个就是爸爸妈妈出去打工，留给他奶奶的，奶奶都 70 多了，也管不了的那个孩子，我们班是怎么也赶不上 Z 老师班了”。这时 C 校长走进来对大家说：“好了，别唠了，各位抓紧时间把各班的成绩拢一下，除了把考试总分核一下，每道题的总分也要核一下，哪些题错了，错的学生多不多，如果错的多了，分析一下咋回事，回头开家长会也好和家长有个交代。记住啊，班级里的学生成绩排名弄好后发给我，除了你们自己知道外，不要发给学生和家长，免得上面查给自己和学校惹麻烦”。C 校长走后，大家都开始忙着做考试质量分析。Z 老师说：“我今年省不少劲，数学考试我班学生全对了，没有错的，不用分析那么多了”；W 老师斜眼瞄了一眼 Z 老师，假装愤愤地说：“你不带这么气人的，我还在想我可咋整，我班学生考的不好，错的那么多”，旁边的 N 老师说：“那你要做的事情可多了”，“可不嘛，我得看看全班学生每道题的得分情况，全对的有哪些？全错的有哪些？错的多的有哪些？错的少的有哪些？还得分析原因”，W 老师嘟囔着说，N 老师开

玩笑说到：“你还得好好找找是不是你自己讲的不好”，P 老师把话接过去：“咱们学校的学生和别的学校的学生太不一样了，想让他们都考好，比登天还难”，“可不嘛”，W 老师附和着。O 老师说：“我觉得考试成绩不好，原因不用分析，把班级考试整体情况在家长会上一说，然后把每个学生的成绩单独发给家长就完事了，干嘛整的那么累”，说完这话 O 老师发现大家都在看着她，她忽然好想意识到什么了：“这是我的心里话，只和你们说，不许传到校长耳朵里啊”大家都会心地一笑。W 老师说：“我今年打算把去年做的考试质量分析里面的改进措施拿过来，教课老师是我、学生还是原来的学生，改进措施也不会变哪去”，“我觉得你分析的对”，“我认为可以”，“我也这么认为的”，“我一直就这么做的”，大家就改进措施问题七嘴八舌地发表了自己的看法，最后大家达成了一致意见。

下午 4 点，C 校长走进大队部，询问考试质量分析进展情况。所有班主任都将自己班级的考试质量分析用手写的简洁式报告提交给了 C 校长，“Z 老师，你们班的数学考试质量分析里面怎么没有错误分析呢？”，“我们班数学这次考试全部满分”。Z 老师洋洋得意地回答说。C 校长又接着翻阅了剩下的考试质量分析报告，抬起头来对着所有班主任说：“你们有没有发现你们做的考试质量分析是有问题的？”，“C 校长，我们都很认真，不会有什么问题吧”，“不可能有问题，我都核对好几遍了”，“有错别字吗？”，班主任们都疑惑地看着 C 校长。“你们的考试质量分析报告都做的头重脚轻：考试得分情况分析挺详细，查找错误的原因和改进措施也太少了”。班主任们松了一口气，几乎异口同声的说“你说的是这个呀”。W 老师站起来笑着对 C 校长说：“C 校长，这不是问题，因为学生和家长主要是看考试分数，关心学生的成绩好坏，所以我们把考试得分情况分析的比较详细，我们认为这也是老师必须要做好的，我们要把考试分数情况在家长会上向家长汇报和交流。原因和改进措施是给我们老师自己看的，我们知道就行了，不用写那么多”，“C 校长，你又不是不知道咱们学校学生比较特殊，留守儿童和外来务工人员子女比较多。平时关注自己孩子的家长又不多，所以，我们就只能用最直接的方式，把考试分数做详细分析并将这些孩子的得分情况汇报给家长，说多了，这些家长也不愿意听。这也是我们为什么特别关注分数情况分析的原因”，H 接着话茬表达了自己的观点，其他老师也附和着点头同意这两位老师的说法。C 校长默认了班主任们的做法，笑着说：“那好吧，你们在家长会上给家长汇报完考试成绩之后，别忘了把自己写的班级考试质量分析报告打成电子版输出一份，然后作为教学材料存档”。

5.小学考试成绩的社会关注

出题、考试、阅卷、考试质量分析的工作都结束了。2019 年 7 月 9 日，B

小学迎来了每学期一次的考试后家长会，由于是第一次出题、阅卷全流程学校自主，因此对于这第一次之后的家长会，C 校长是非常重视的。然而家长会当天，在对各班进行巡视时，她发现，原本一个 32 人的班级，只来了 5 名家长。这个班级的班主任 F 老师对 C 校长说道：“我们班学生群体中特殊儿童多，爷爷奶奶看管的留守儿童和外来务工人员子女占比都很大，爷爷奶奶岁数大了，不来开家长会也在情理之中”。说话间，F 老师脸上的无奈早已显露无疑。

B 小学的学生考试成绩和往年没有太大差异，这次小学考试成绩可能又一次决定了 B 小学在全市所有小学中的排名，C 校长隐约感觉到全市教育大会上又因为排名倒数而被点名的危险：“D 校长，今年的全市教育会你去参加吧？”，“为什么？”，D 校长不解地问道，“我心理承受不了了”，C 校长半开玩笑地回答。“哈哈哈”，两个人都大笑起来。

又到了全市的教育总结大会，今年虽然进行了改革，但是 C 校长还是被点名批评的那个，对于这样的“被点名”，C 校长似乎早已经习以为常。会议结束后，在会场门口遇见了 F 区的 Q 校长，Q 校长与 C 校长是原来在中师的同学，Q 校长煞有其事地把 C 校长拉到一边，和 C 校长说：“现在是我们自己出题了，你想点办法搞搞成绩，这对于咱们校长来说不是最轻而易举的事儿嘛”。说着还冲 C 校长眨眨眼，C 校长笑了笑，没有说什么，Q 校长反趣道：“你这个人啊，就是脑筋不够用！”。回到学校后，C 校长找到 D 校长，和 D 校长讲了别的学校提高成绩的“有效办法”：“在我们小学这个圈子里，有一些提高学生成绩的方法，被大家普遍接受的。例如，有的学校在考试之前，将出完的题，领着学生做了好几遍，正式考试时能够打满分；有的学校故意将试题难度降低，出一些大众题，学生都得高分，让家长和学生都高高兴兴的”。D 校长说：“虽然我们学校不如特别好的学校那么备受社会关注，但也绝不会在考试环节上弄虚作假。我们学校现在不这么做，将来也不这么做，我们必须要为学生负责”。C 校长着急地说：“道理我都懂，我也有教师道德和良心。但是我们学校怎么做才能让学生成绩提高，不再因成绩不好年年被点名呀？”，D 校长看看 C 校长：“这需要我们探索和研究，我们一起努力吧，愿我们的学生将来都能很好地在社会上生活！”。

结语

B 小学的考试从市里统考到区里统考，再到学校自己负责考试，呈现出编制试题的权力逐渐下放的趋势，但是依据教师用书和教材为依据编制试题、试卷的评阅程序和环节、考试质量分析等环节并没有太大变化。2016 年我国提出核心素养培养问题。小学阶段是基础教育中的重要阶段，目前小学教育中核心素养培养效果怎么样？小学考试是否是针对核心素养的培养目标设计与实施的？作为小学教育专业的教育硕士研究生有必要从理论和实践中明确一个问题：基于核心

素养培养的要求，小学考试与评价改革需要我们作为教育工作者做些什么。

教学指导手册

小学考试的那些事——基于核心素养的考试与评价改革

1.教学目标

教育硕士借助先前的课程与教学论知识、测验的编制与实施、测验项目的质量分析、标准化测验的基本程序与应用等对教育测量的质量指标中效度问题形成明晰的认识，初步建构教育评价与测量中效度内容的知识结构体系。

具备教育测量的效度质量指标分析能力：能够基于考试实施的各个环节，掌握判断考试有效性高低的指标及其影响因素；学会估计并提高小学考试效度的能力。具备试卷编制与质量分析报告应用能力：能够基于核心素养要求编制高质量试卷；将小学考试质量分析报告用于指导小学教师的教育教学工作、学生的学习以及学校管理等工作。

1.1 适用课程

本案例主要适用于课程《教育评价与测量》中测验的编制与实施、测验项目的质量分析、标准化测验的基本程序等相关知识的学习。

1.2 教学对象

本案例主要为小学教育专业领域硕士研究生开发，也适用于小学教育专业本科生。

1.3 具体教学目标

（1）知识和能力层面，能够对教育评价与测量中考试效度的内容深入理解，掌握考试效度的关键能力。

（2）认识和情感层面，能够聚焦教育评价与测量中的考试效度并形成正确认知，体验考试效度在小学教学中培养学生核心素养的价值。

（3）分析与综合层面，能够通过考试有效性的现象揭示考试效度对小学教育教学的导向功能，从整体上把握效度在教育评价与测量知识结构中的重要性。

（4）创新与应用层面，能够在现实和理论的基础上创新性地将核心素养培养与小学教学关系通过考试的方式反映出来，从而有效的指导小学教育教学工作。

2.启发思考题

（1）小学考试与评价的基本流程是怎样的？

（2）小学考试效度的判断标准是什么？

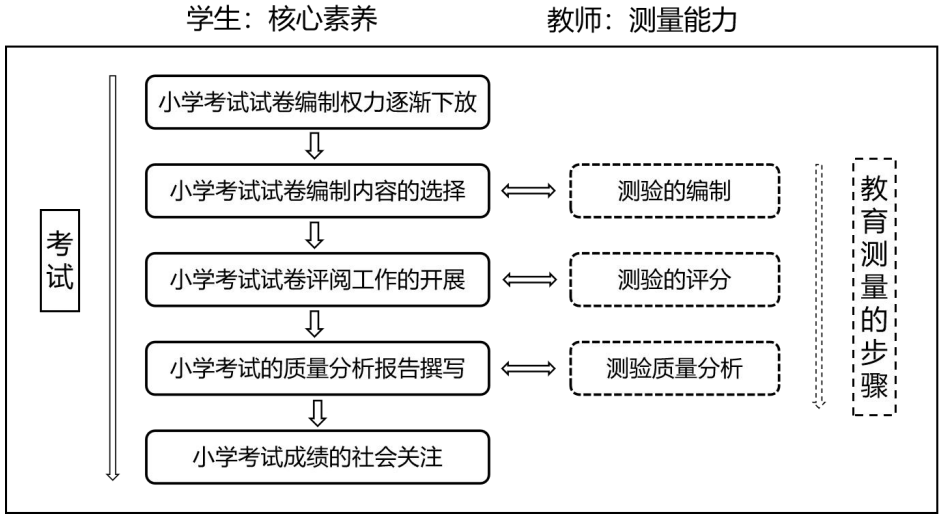
（3）小学考试效度的影响因素有哪些？

（4）如何确保基于核心素养培养目标的小学考试有效性？

（5）核心素养需要小学考试在哪些方面改革？

3.分析思路

首先，按照考试的运行程序对“小学考试试卷编制权力逐渐下放”、“小学考试试卷编制内容的选择”、“小学考试试卷评阅工作的开展”、“小学考试的质量分析报告撰写、”“小学考试成绩的社会关注”五个问题进行了剖析与讨论。其中，“小学考试试卷编制权力逐渐下放”呈现了从市级行政管理部门到区里行政管理部门再到学校管理出题权力下放的过程。“小学考试试卷编制内容的选择”呈现了小学考试的试卷内容选择依据。“小学考试试卷评阅工作的开展”呈现了小学考试的评阅者及其评阅程序。“小学考试的质量分析报告撰写”呈现了小学考试质量分析的用途。“小学考试成绩的社会关注”呈现了小学学生考试成绩的高低带来的影响。其次，在现实状况呈现的基础上，分析小学出题权力下放的背景，讨论试题编制缺少与核心素养结合的原因、试卷评阅结果能够客观公正的途径、考试质量分析用于指导教师教育教学和教学管理中的缺位，教育要求与社会认可之间的学校承受的冲突压力，等问题。这些问题的讨论与分析主要是澄清教育评价与测量课程中效度的整体知识结构及其价值。最后，在讨论与分析基础上，师生共同探究小学考试的有效性，即，考试效度作为衡量教育目标实现程度的有效提升的路径与策略，为将来从事小学教育教学工作指明方向。



4.案例分析

4.1 案例回顾

本案例选自 2019 年 10 月，M 市 A 区 B 小学的 C 校长对小学考试问题的访谈

与调研。B 小学的考试从市里统考到区里统考，再到学校自己负责考试，呈现出编制试题的权力逐渐下放的趋势，但是依据教师用书和教材为依据编制试题、试卷的评阅程序和环节、考试质量分析等环节并没有太大变化。2016 年我国提出核心素养培养问题。对于小学阶段学生的核心素养如何培养？如何针对核心素养的培养进行试卷的编制、实施与评价？考试后的试卷质量分析如何能发挥最大效力？是小学阶段教师领导需要深入思考的问题。

4.2 理论基础 1——编制测验双向细目表

测验双向细目表是测验的蓝图，具有极大的实用性，对教学和测验都有重要作用。它是教育目标、测验目标的具体化，是测验编制的重要依据，可以帮助出题者明确教学内容是什么、如何支配教学时间，明确相应知识内容的目的是什么以及如何用测验来测量相应教学内容的目标实现程度等。

双向细目表的编制需要按照相应的步骤来进行：

（1）列出教材大纲；

（2）对列出的教材大纲进行权重；

（3）对各种教育目标进行权重；

（4）编制双向细目表。从内容和目标两个维度编制测验的双向细目表。

4.3 理论基础 2——主观性试题的相对评分法

相对评分法是根据对学生答卷之间的相互比较进行的评分的。最好的答案得最高分，最差的答案得最低分。此种方法适用于测量较为复杂的行为，且难于确定标准答案的时候。

使用此种方法，第一步，确定评分等级并确定每一等级应占的比例。评分等级一般可采用 0-3，0-5，0-10。等级分的太细，分数跨度太大（0-20），不仅给评分带来困难，而且会把实际是同一水平的答卷在那些并不重要的方面加以区分。各等级所占比例，可视学生的实际水平而定，也可按正态分布的理论加以划分。第二步，粗阅答卷，如把试卷分为一般、高于一般、低于一般三大类。第三步，细读答卷，检查粗读分组是否精确，并根据确定的等级数量及比例对答卷分类，评定分数。

4.4 理论基础 3——测验质量分析的作用

（1）测验质量有助于测验科学化的实现

有助于编制测验的技能、技巧的提高，能为进一步修改试题提供依据，并且有助于提高测验的信度、效度，经测验质量分析鉴定出的达到各项指标要求的高质量试题，可保留起来，以备后来搭配试卷使用。以后的测验就可以根据已确定的测验目的，从题库中抽取试题，拼配试题。可以省时、省力，可以保证测验的

高质量，大大提高测验的科学性。

（2）测验质量分析有助于提高教学质量

通过对测验结果的深入分析，才能发现教学中存在的问题，起到改进教学的作用。通过测验分析，既可以发现学生学习上的共同弱点，也可以诊断个别学生在学习上的症结所在。有助于测验结果的合理解释，学生才会最大限度地利用测验所提供的信息，来促进自己的学习。

4.5 理论基础 4——核心素养

2016 年 9 月 13 日，中国学生发展核心素养研究成果大会在北京召开。会上敲定核心素养三方面六素养十八大要点。对核心素养三方面进行分析的基础上发现，中国学生核心素养实质上呈现 A-BCD 模式，即以培养“全面发展的人”（All-round development）为核心，分化为文化基础（Basis of culture）、社会参与（Community participation）、自主发展（Develop independently）三个方面。在案例中，进行试卷的编制时，除了考量学生对于书本知识的掌握情况外，还需在不同试题的编制时关注学生几方面核心素养能力的提升。

5.课堂设计

5.1 时间安排

大学标准课 4 节：180 分钟。

5.2 环节安排

（1）布置预习并分组讨论。用 40 分钟时间，阅读材料并围绕思考题进行组内讨论；各组将讨论的结果总结并提炼出来；

（2）组间交流。用 50 分钟时间，每组派一名代表结合思考题将自己组的结论进行分享，阐述自己的组的观点，每道题约 10 分钟；

（3）讨论各组生成的问题。用 20 分钟时间，全班共同参与探究通过阅读和讨论案例过程中生成的问题。

（4）教师点评。用 20 分钟时间，对讨论进行概括与进一步引导，明确共识与分歧点。

（5）设计思维导图：用 20 分钟时间，设计教育评价与测量的知识结构图谱。

（6）头脑风暴。用 30 分钟时间，全班进行头脑风暴，思考当前小学考试效度作为衡量教育目标实现程度的有效提升的路径与策略，以及自己将来从事小学教育教学工作的方向。

5.3 人数要求

30 人以下的班级教学。

5.4 教学方法

讨论、合作、参与式教学；小组与班级讨论为主；教师点评为辅。

5.5 活动设计建议

（1）提前三周布置阅读任务。包括案例文本、教育评价与测量教材和理论文献。

（2）设计与安排学生的小组讨论。选择适合分组讨论的教室上课；每个小组指派一名学生作为负责人，负责准备讨论与记录的材料，并负责小组讨论过程中发言人的主要观点与结论的记录工作。

（3）教师准备好点评的资料和提纲。对学生小组的讨论进行点评，适时的提升理论，把握教学的整体进程。

（4）指导学生在课下对自己设计的知识结构思维导图和头脑风暴中的路径与策略进行完善，教师给予必要的指导。

（5）教师课后根据各小组的讨论和汇报情况，以及生成问题的质量及时分析案例教学的得失，以便为今后的案例教学做进一步的改进与完善。

6.要点汇总

编制试题的权力逐渐下放，学校领导教师需要具备教育测量的相关能力素养，结合核心素养对学生的基本要求，体现在试卷编制前、编制过程中，同时在测验结束后评分阶段，依据相关标准进行客观评分，后续结合学生成绩进行试卷质量分析，为后续标准化题库的建立奠定基础。

6.1 测验目的的确定

要编制测验首先要明确测验目的，测验目的明确规定测验所要达到的预期结果或标准，是编制测验的出发点和依据，影响着测验的质量高低。只有明确测验目的，我们才知道测什么、怎么测，才能解决一系列关系到测量质量优劣的重要问题。测验的质量在一定程度上是由测验的效度决定的。如果缺乏明确目的的测验不仅是盲目的、无效的，而且是有害的。对于测验目的的确定，要结合学生核心素养的几个方面，将核心素养落实到后续测验的编制中。

6.2 测验试题的编制

确定了测验的目的，编制出测验双向细目表，接下来的工作就是编制试题，即所说的命题。命题是测验编制中的核心环节。它包括选择试题类型，编拟试题，确定评分方法，编制测验说明，配搭、组织试卷等项工作。

根据不同的标准，可以把试题分为不同的类型。编制试题的第一步工作，便是选择适宜类型的试题。试题分类的依据很多，最常用的方法是依据试题评分是

否客观，把试题分为“客观性试题”“非客观性试题”两大类。

6.3 试题的评分

评分是教育测量的重要环节，对测验质量有着极重要的影响。评分应力求准确、客观，只有这样，所评定的分数才是可靠的。

6.4 试卷的质量分析

对测验质量进行分析，有两方面的作用：一是可以为筛选、修改试题提供依据，提高测验质量，使测验逐渐实现科学化。二是可以为改进教学、提高教育质量服务。

7.推荐阅读

[1]马婷.表现性评价的设计与实施——基于浙江省小学考试评价改革的实践案例基础[J].教育课程，2018（11）.

[2]蔡雪萍.素质教育背景下小学考试改革的几点思考[J].新教师,2016(9).

[3]施久铭.“一师一卷”：用“考试”打败考试——一所小学从评价改革到核心素养的实践探索[J].人民教育，2014（11）.

[4]李协京.从日本全国学力考试看其中小学教育质量监测[J].外国教育研究，2008（10）.

[5]方超，黄斌.中国过度教育测量、趋势及其影响因素的实证研究[J],教育科学，2018（8）.