

中考语文阅读开放题与 PISA 阅读开放题例析

文 申宣成

〔摘要〕本文以 PISA2009 阅读试题、2009 年上海中考语文试题为样本,从试题数量、问题和指导语、试题难度和评分标准等方面比较了 PISA 阅读开放题和中考语文现代文阅读开放题的异同,为中考语文开放题的设计提供了参考。

〔关键词〕PISA 中考语文 阅读 开放题

自从教育部在《关于 2000 年初中学毕业、升学考试指导意见》中明确提出“应设计一些开放性试题,鼓励学生有自己的见解”之后,我国的语文中考日益关注开放性试题。一般来说,多数人倾向于把那些答案灵活开放(即可以有多个正确的答案)的题目都归入开放题^①。相比国内而言,国际学生评估项目(Programme for International Student Assessment,简称 PISA)命题专家对开放题的界定更为明确:开放题(Open-constructed response,全称为开放结构式反应题)是指那些答案灵活、应答内容较多、通常要求对答案做出解释的题目^②。而那些答案虽然也较灵活,但不需要做出较多的解释,通常只需要用一个词或短语来回答的题目,称为封闭结构式反应题(Closed-constructed response)。按照 PISA 的定义,中考语文中诗词默写类试题以及部分仿写改写类试题就不属于开放题的范畴。

2009 年 4 月,上海市正式参加了 PISA2009 测试,这为进行 PISA 和中考的国际比较提供了第一手的材料。下面以 PISA2000 阅读开放题和上海 2009 年中考语文试运转卷现代文阅读的开放题为样本,从试题数量、问题和指导语设计、试题难度和评分标准等几方面作一些分析^③。

一、试题数量的比较

PISA2000 阅读测试和 2009 年上海中考语文现

代文阅读的题型和数量如下表:

表 1 PISA2000 阅读和 2009 年上海中考语文现代文阅读题型数量比较

数 量 类 型	标准多 项选择	复杂多 项选择	封闭结 构式应 答题 (填空题)	开放结 构式应 答题 (开放题)	合计
PISA2000 阅读	56	7	15	63	141
2009 年上海中 考语文(现代 文阅读)	2	0	7	2	11

从上表中可以看出,在 PISA2000 提供的 141 个题目中,开放题就占了 63 个,约占试题总数的 45%,是数量最多的题型(141 个题目被分成 9 组,每组题目按 30 分钟设题,每个被测的学生需要做其中的 2~3 组阅读题目^④,因此,在同一次测试中,不同学生所做的试题可能是不同的)。而上海的中考现代文阅读共有 11 个题目,开放题占了 2 个,尚不足 20%。两者的比例相差如此之大,原因在于评价旨趣的不同。PISA 阅读评价的目的是考察学生理解、运用、反思和评价文本的能力,这些能力更适于使用开放题;而中考在重点考察学生理解和运用语言的能力的同时,还要兼顾学生对学科知识的记忆和理解,如字词、修辞、句式等。《上海市 2009 年初中学业考试手册》在语文科现代文阅读能力测试方面就列出了 13 个考察要点。

二、问题和指导语设计的比较

问题和指导语又称为题干。命题者通过题干明确答题的要求,落实评价的目标;被测试者则要根据题干明确答题的任务。问题和指导语设计得清楚与否,直接影响到测试的效度^[5]。

2009年上海市中考现代文阅读设计了两个开放题。第一个是针对文本《漫谈桂文化》的最后一句话设计的,其表述如下:

你对文末提到的专家的建议(画线部分)有什么看法?请简述理由。(5分)

第二个是针对文本《月光启蒙》的最后一句话设计的,其表述如下:

如果以“母亲失去了记忆,而我心中永远珍藏着那一份母爱”来结尾,和原文比较,哪个好?为什么?(80字左右)(8分)

这两个开放题对问题的表述明确具体,不易产生歧义;设题紧扣文本,一个侧重思想内容,一个侧重表达形式,可谓开合有度,各有侧重。

但是从其指导语的设计来看,也存在一些问题,如第2个开放题的指导语中“80字左右”就需要斟酌。命题者的本意可能是担心学生的答案写得过多,以此提示考生要简单陈述,不要罗嗦,可是这一要求同时有可能暗示学生要答到80个字,不然就达不到要求,这其实是一种误导,因此用“不超过80个字”表述更合适。2008年的上海高考语文试卷的第5题也有“80字左右”的要求,而2009年的上海语文高考试卷的第6题,就改为“不超过45个字”。

相比之下,PISA阅读的开放题更注重设计指导语。以下是PISA2009样题中的一个问题的设计语:

题14:涂鸦

我们可以讨论一封信件叙述的事情(它的内容)。

我们可以讨论一封信件写作的手法(它的风格)。

不论你同意哪个作者的论点,你认为哪一封信写得比较好?请根据其中一封信或者两封信的写作手法来解释作答。

试题设计者在问题提出之前,设计了两句指导语,明确了叙述内容和叙述风格的区别,解释了问题中“写作手法”这个关键术语,这就避免了学生因为对这个术语不理解而影响答题,这一指导语充分体现了PISA阅读测试的特点:目的不是考察学生的语文学科知识,而是考察学生解决问题的能力。

三、试题难度设计的比较

试题难度的把握,则主要涉及到测量信度的问题。PISA试题都经过了大规模的试测和精确的数学分析,“打造的是把钢性的量尺,试题难度不会像传统测试的弹性量尺随受试人群样本的能力不同而变化”^[6]。如上面引用的样题中的第14题,其难度水平为4级,对应的分数为581分。在PISA的试题地图(PISA Item Map)和阅读素养水平地图中,这一水平和分数表明学生可以“运用常识或假设,批判性地评价文本,展现的是对篇幅较长或内容较为复杂的文本的精确理解”^[7]。试题难度的准确定位为国际间学生素养的横向和纵向的比较提供了科学的依据。除此之外,PISA还在每次测试时使用一些相同的题目,这些题目称为“锚题”(Anchor Item),以便对不同年度的测试结果进行更为准确的比较。

与PISA相比,我国中考在试题难度的确定方面采用的是传统测试理论,试题的难度没有准确的定位。尽管2009年上海中考的考试细则中规定:“试题的难易比例约为1:1:8”,但这仅仅是一个粗略的估计,具体到某一份试卷、某一个题目,并没有足够的证据定位其难度等级。以上文提到的2009年上海市中考的两道开放题为例,很难说哪一个难度更大一些;而分值一个为5分,另一个为8分,也很难说出什么依据。可能的解释是参考答案中两个题目的要点多少不同,而这样给试题赋分恐怕并不合适,因为决定一个题目分值的,首先是完成题目所需的文本阅读的深度和长度,而不是答案要点的多少。

中考采用传统的测量方式,优点在于简单易行,试题开发的技术要求和成本较低,但是不利于测试的横向和纵向比较,不利于质量的分析和监控。从长远考虑,中考应该借鉴PISA测试的优点,依托专业命题机构,准确定位试题的难度,建立自己的试题库。

四、评分标准的比较

语文开放题的优势是便于学生表达自己的主观感悟、张扬自己的个性特点、展示较高层次的认知能力,但是它却给评分工作带来了难度。如何使开放题的评分尽可能客观公正,是PISA和中考面对的共同挑战。那么,两者在这一问题上的处理有什么不同呢?这里以2009年上海市中考语文的第2个开放题(在整个试卷中的序号为第26题)和PISA样题的

第14题提供的评分标准为例略作分析。

上海市中考第2个开放题的评分标准:

(8分)观点(1分)理由(6分)(写出两点,每点3分)语言(1分)

原句好的理由:(1)意境更含蓄(2)结构呼应更紧密(3)内容上突出了诗情(4)语言上具有诗歌的韵味

改句好的理由:(1)直接点明母爱主旨,更加鲜明地表达了对母亲的爱(2)呼应开头

PISA 样题第14题的评分标准:

满分:代号1

就其中一封信或全部两封信的风格或形式,解释意见。

可以指出下列特点:

写作风格、论证的结构、论证是否中肯、语气、语体(正式或非正式)及说服读者的策略。用词,例如“较好的论点”等必须有证据支持(注意:“有趣”“易于阅读”和“清晰”等字眼不够具体)

●嘉嘉的较好。她提出不同的观点让你思考,她也提及涂鸦艺术家对环境的破坏,我认为这点很重要

●嘉嘉的文章较有效,因为她直接提到了涂鸦艺术者。

●我认为两封信中,嘉嘉的信件较好。我觉得小雅有点偏见。

●我觉得小雅提出了很有力的论点,但嘉嘉的结构较好。

●小雅的较好,因为她没有针对任何人(以内容的特点去解释自己的选择。若诠释为“没有攻击任何人”,则是一个更清晰的解释)

●我喜欢嘉嘉的信件,她有力地表达意见。

零分:代号0

就同意或不同意作者的立场作出判断,或只简单地意译原文或评论文章内容。

●嘉嘉。我同意她的所有说法。

●嘉嘉的信件较好,正如她所说,涂鸦既花钱又浪费。

●小雅的信件较好,她所说的都重要。

或:判断欠充分解释。

●小雅的信件是最好的

●小雅的信件较易阅读

●嘉嘉的论点较佳

或:答案不合理或与问题无关。

●嘉嘉的写法较佳,她一步一步分析问题,以此为基础,她最终得出一个合理的结论

●小雅的信件较好,因为她从头至尾都坚持自己的立场。

没有作答:代号9

从上面的两个评分标准不难发现:

上海市中考开放题的评分标准是累积式记分的,

即把总分分解为若干个得分点,完成一个得分点就可以得到相应的分数,它的优点是方便灵活,便于操作。而其缺陷在于:观点分(1分)和语言分(1分)考察的都不是学生的鉴赏能力。前者几乎是白送的,因为开放题的观点本身就不存在对与不对的问题,只有理由是否充分的问题;后者测量的是语言表达能力,而不是对内容的理解程度,甚至可能误导评分者片面关注学生的语言表述。这些都影响了评价的效率,即所评的东西不是自己想要的东西。

PISA 的评分标准是水平式记分的,即只要达到了测量的水平,就可以得满分或部分得分。在题14所罗列的满分答案中,例1显然比例2要完善得多,但是两者都是满分,因为例2同样达到了等级4这一水平;而且,该标准集中在鉴赏能力的评价上,而不考虑表达是否通顺、书写是否规范等其它因素,这体现了PISA对试题效度的要求:只测量自己想要的东西。为了便于评分者对答案做出准确的判断,PISA开放题在评分标准中收录了试测过程中学生的典型答案(如题14中罗列了6个满分答案和8个零分答案)。此外,PISA还对学生的答案分类编码,这为考试结果的统计分析提供了很大的便利。

参考文献:

[1]参见姚永夫.开放性试题的特点及其解法初探[J].中学文科,2001,(8).谢进军.中考语文开放试题比例逐年增多[N].中国教育报,2006-7-29:7.

[2]Ray Adams and Margaret Wu 编, PISA 2000 Technical Report, p21 <http://www.pisa.oecd.org/dataoecd/53/19/33688233.pdf> 浏览日期:2009年7月14日。

[3]因为PISA2009的评价报告2010年才能公布,所以本文采用了PISA2000阅读评价报告《为了改变的阅读》(Reading For Change)中的数据;而样题则是从PISA2009的样题中选取的。采用2009年上海中考语文试运转卷作为分析样本是因为在笔者看来,仅就开放题的设计而言,它比该年度的正式卷更有典型性。

[4]Ray Adams and Margaret Wu 编, PISA 2000 Technical Report p22 <http://www.pisa.oecd.org/dataoecd/53/19/33688233.pdf> 浏览日期:2009年7月14日。

[5]方云龙曾做过详细的举例分析,见《开放性试题质量分析》,《语文教学通讯·初中刊》2002年第9期。

[6]王蕾.Rash 客观等距测量在PISA中国试测研究中的实践[J].心理学探新,2007,(4):73.

[7]Irwin Kirsch 等,Reading for Change, p25,<http://www.oecd.org/dataoecd/43/54/33690904.pdf> 浏览日期:2009年5月19日。

[中宣成 华东师范大学课程与教学研究所 200062]