

中考物理试题的因子结构分析及其启示*

翟彦芳¹ 邢红军²

(1. 北京师范大学中国基础教育质量监测协同创新中心, 北京 100875; 2. 首都师范大学教师教育学院, 北京 100048)

摘要:基于义务教育物理课程标准与考试大纲,通过对北京市高级中等学校招生考试(中考)物理试卷的真实数据进行因子分析,寻找试卷的3因子结构,并进一步建构了不同的模型。基于不同模型的比较并结合理论分析,寻找最优模型。揭示了中考物理试卷的真实结构,这为中考物理命题提供了有益的启示。

关键词:初中物理;因子分析;中考试题

中图分类号:G634

DOI:10.19789/j.1004-9398.2020.05.014

0 引言

世界各国都将提高学生的科学素养作为科学教育的核心目标^[1-3]。例如,美国《新一代科学教育标准》(next generation science standard, NGSS)^[4]、国际学生学业成就评价(program for international student assessment, PISA)项目^[5]及美国国家教育进展评估(national assessment of educational progress, NAEP)项目^[6]。在我国,作为义务教育阶段大型学业测评的高级中等学校招生考试(中考),物理是重要的考察学科之一,中考物理试题能否全面地反映学生物理知识和技能的掌握情况,就成为一个值得探讨的问题。相关研究大多聚焦在中考物理试题的内容一致性分析上^[7-8],试题中某个专项题目的偏质性研究^[9-13],研究多采用试卷分析法和经验总结法进行质性分析及一般信度、效度分析,鲜有研究关注中考物理试题的结构。因此,本文在综述国际大型测评项目的基础上,以北京市近年中考物理试题为例,探究中考物理试题的因子结构,以期中学物理教育教学改革及评价提供有益的启示。

1 中学生学业成就测评框架比较分析

PISA项目从2000年开始对科学素养进行测评,2006、2015年主要监测的是科学素养领域。PISA 2000借鉴了Miller^[14]的3维度模型,经过PISA

2006对科学素养进行了重新定义之后^[15],PISA 2015明确了科学素养是一种能力,提出科学素养是指:“处理与科学相关事务的能力,对科学和技术作出合理解释,成为有科学思想的批判思维的公民”^[16]。这一定义包含了3种能力:科学地解释现象的能力,评价并设计科学探究问题的能力与科学地解释数据和证据的能力。PISA科学框架涉及的主要内容领域包括:物理科学、生命科学、地球与空间科学、科学技术及互相交叉的一些内容^[17]。

国际数学和科学评测(third international mathematics and science study, TIMSS)项目强调对科学学科评价不能脱离背景,要在一定的内容领域下进行评价,并且要通过认知领域的划分及描述说明一系列的思维过程^[18]。其中测评的内容领域有:生命科学、自然科学、地球科学、生物学、物理学、化学和地理科学。认知领域有:了解、应用和推理。此外,将科学探究融入对科学内容和技能的掌握和运用中,要求学生掌握科学的知识和方法,并掌握提出问题及假设、设计实验、呈现数据、分析和解释数据及证据,得出结论并做出解释等一系列过程及能力。

NAEP项目(以NAEP 2011为例)对科学监测的内容领域包括^[19]:物理科学、生命科学、地球和空间科学。科学实践能力包括:①识别科学原理的能力。重点培养学生在3个知识领域中的每个基本原理的了解、认识、记忆、定义和描述的能力。②运用科学原理的能力。重点培养学生利用科学知识准确地解释和预测自然界中观察到现象的能力。③运用科学探究的能力。重点培养学生设计、批判性思维及评价科学探究,能够采取恰当的数据类型,并论

收稿日期:2018-11-26

*北京市教委“十三五”规划社科基金重点课题(SZ20171002815)

证结果的能力.④运用技术设计的能力.重点培养学生在系统中运用科学的知识和技能,为解决现实世界中的实际问题提供建议的能力.

通过比较以上3个国际测评项目中的科学素养测评框架,均包含内容领域、能力/认知领域.内容领域均包含物理科学及物质科学相关的主题,认知领域则注重对科学原理的理解、分析,对现象的解释和科学探究的能力.并且TIMSS和NAEP项目均提出将科学实践融入到对内容掌握的过程中,通过科学探究让学生自己构建科学的知识体系,生成对科学的理解.但是,在测试内容、范围和具体指标上有较大不同:①测试内容范围不同.PISA重在培养学生应对社会的能力,故在测试内容上不限于课堂所学知识,TIMSS着重考察各国课程中的共同内容,故测试内容是各国科学教育内容的交集.②测试框架重点不同.PISA重在能力指标,TIMSS重在内容指标,NAEP重在科学实践.③测试指标不同.PISA不仅从内容、认知维度进行了测查,还从科学情境、科学态度全方位对科学素养做出了评价指标.

对我国初中物理课程标准(以下简称课标)开展分析,梳理物理课程的测试框架,物理测试主要从内容维度、认知维度和情感态度价值观维度进行考察^[20].其中内容维度划分为科学内容(包括物质、运动和相互作用)以及科学探究;认知维度则按照了解、认识和理解3个维度进行划分^[21-22];情感态度价值观维度则考察学生科学态度的养成,对科学的兴趣及求知欲等.

通过与国际大型监测项目的科学测试框架比较显示:①科学测试内容一致.课标中规定的物理内容领域涵盖了TIMSS和NAEP项目中涉及的3个内容领域.②科学探究的测试方式一致.与TIMSS项目类似,课标认为科学探究应体现在教学过程中,并且融合在科学知识及技能中综合考察.③课标包含了情感态度价值观的测查.与PISA项目类似,优于TIMSS和NAEP项目,规定了在情感态度价值观方面科学素养的目标.

2 中考物理试题的因子结构分析

本文基于北京市某区教育考试中心提供的2014年该区中考物理学科的学生作答反应数据,将该区学生在2014年北京市中考物理试卷的作答反应作为研究对象,按照50%的比例随机抽取样本2238例进行研究.研究程序如下:

(1)了解试卷概况,对试卷的题型、难度和区分度进行分析.

(2)因子结构分析.从不同角度构建试卷评价模型,通过模型拟合指数的比较对模型进行比较评价,从而总结出中考物理试卷中潜在的结构.其中,样本学生随机分为2组,一组学生用于探索性因素分析,寻找可能潜在的因子,构建潜在因子结构模型;另一组学生用于验证性因素分析,检验各个模型的拟合水平.各因子结构分类如下:1)单因子结构:物理能力.从整体上看,试卷总体测试的是物理相关的内容及能力.因此,首先将所有题目归属于1个因子,即物理能力的单维模型.2)内容3个因子结构:物质、运动和相互作用及能量.依据课标中内容维度的划分,中考物理试卷覆盖的内容分为物质、运动和相互作用及能量3部分.3)认知2个因子结构:了解和理解.结合课标中对认知水平的划分以及2014年北京市中考物理考试说明中提供的双向细目表,将认知水平划分为“了解”“理解”2个层次,即构建认知维度的2个因子结构模型.4)可能的因子结构.通过探索性因素分析寻找试卷中潜在的因子结构,构建新模型.基于理论研究,并结合模型1~3比较结果以及探索结构模型4的结果,构建新模型.同时,比较模型得出最优模型.

(3)根据模型比较结果,对中考物理考试反映出学生对课标规定的内容、能力、掌握程度以及试题真实的结构展开讨论与分析,对今后中学物理教育教学实践与改革提出意见和建议.

本文采用SPSS 20.0和Mplus7软件进行数据分析.采用以下指标及评价标准分析模型与数据拟合效果:拟合优度卡方检验,卡方(χ^2)越小越好;自由度(d_f); χ^2 与自由度之比(χ^2/d_f) <3 ,拟合较好;近似误差均方根(root-mean-square error of approximation, RMSEA) <0.050 ,越小越好;比较拟合指数(comparative fit index, CFI) >0.900 ,越接近1越好;Tucker-Lewis指数(TLI) >0.900 ,越接近1越好;加权均方根残差(weighted root mean square residual, WRMR) <1.000 ,越小越好. $P < 0.05$ 为差异有统计学意义.

2.1 试卷概况

从试卷题型、难度和题目区分度3个方面描述试卷概况:

(1)题目数量为38道题.题目类型有选择题、填空题、实验与探究题和计算题,对应分值分别为:

40、12、35 和 13 分. 总分为 100 分, 作答时间为 120 min. 根据学生作答计分点特征, 分析用题目中的第 29 ~ 32 和 35 道题均按照小题计分, 本次研究将其作单题处理. 故本次研究中题目数量为 46 道题.

(2) 试卷题目中, 较易难度、中等难度和较难题目分别有 33、9 和 4 道题, 试卷难度分布基本合理.

(3) 试卷题目区分度在 0.3 ~ 0.4 有 1 道题, 其他题目区分度均在 0.4 以上, 试卷题目区分度良好. 试卷质量较高, 内部一致性信度达到 0.937, > 0.800. 通过降维分析, 可知该试卷有 4 个主要成分, 这 4 个主要成分解释的方差占了总方差的 50.1%, 该试卷具有良好的信效度.

2.2 模型构建及比较

从不同角度构建模型评价试卷, 通过比较拟合

指数对模型进行评价, 以探索试卷中潜在的内部结构. 包括基于整体、认知和内容维度构建模型的比较及基于因子探索结果构建新模型的比较.

2.2.1 基于整体、认知和内容维度构建模型比较

从整体上看, 试卷总体测试的是物理学科相关的内容, 首先构建物理能力的单维模型 1; 其次根据课标中内容维度的划分, 试卷覆盖物质、运动和相互作用及能量, 基于此构建内容 3 个因子结构模型 2; 最后依据课标构建包含了解、理解 2 个因子的认知结构模型 3.

模型 1 ~ 3 比较. 通过对模型 1、2 和 3 的验证性因素结果指标分析(表 1), 3 个模型的 CFI(0.987) > 0.900, TLI (0.986 ~ 0.987) 均 > 0.900, RMSEA (0.028 ~ 0.029) 均 < 0.050, WRMR (1.152 ~ 1.171) 均 > 1.000, 模型拟合结果均较差.

表 1 所有模型验证性因素分析

模型	因子	题目载荷	χ^2	d_f	χ^2/d_f	CFI	TLI	RMSEA	WRMR
1	物理能力	0.618 ~ 0.969	1 915.2 *	989	1.94	0.987	0.986	0.029	1.171
2	物质	0.791 ~ 0.963	1 911.0 *	986	1.94	0.987	0.986	0.029	1.169
	运动和相互作用	0.459 ~ 0.939	—	—	—	—	—	—	—
	能量	0.646 ~ 0.910	—	—	—	—	—	—	—
3	了解	0.677 ~ 0.960	1 874.3 *	987	1.90	0.987	0.987	0.028	1.152
	理解	0.472 ~ 0.959	—	—	—	—	—	—	—
4	科学知识与方法	0.679 ~ 0.970	1 610.6 *	986	1.63	0.991	0.991	0.024	1.029
	难度	0.731 ~ 0.976	—	—	—	—	—	—	—
	科学应用	0.478 ~ 0.934	—	—	—	—	—	—	—
5	物质	0.928 ~ 0.971	1 834.0 *	983	1.87	0.988	0.987	0.028	1.133
	运动和相互作用	0.460 ~ 0.942	—	—	—	—	—	—	—
	能量	0.648 ~ 0.913	—	—	—	—	—	—	—
	难度	0.728 ~ 0.981	—	—	—	—	—	—	—
6	了解	0.701 ~ 0.970	1 791.6 *	986	1.82	0.988	0.988	0.027	1.113
	理解	0.472 ~ 0.957	—	—	—	—	—	—	—
	难度	0.729 ~ 0.981	—	—	—	—	—	—	—
7	科学知识与方法	0.681 ~ 0.970	1 713.7 *	988	1.73	0.990	0.989	0.025	1.081
	科学应用	0.478 ~ 0.938	—	—	—	—	—	—	—

注: —指无数据; * $P < 0.05$; CFI 为比较拟合指数; TLI 为 Tucker-Lewis 指数; RMSEA 为近似误差均方根; WRMR 为加权均方根残差.

2.2.2 构建新模型及模型比较

通过探索性因素分析试卷中潜在的因子结构, 构建可能的因子结构模型 4. 基于理论研究, 并结合模型 1 ~ 3 及探索结构模型, 构建新模型 5 ~ 7 并进行模型比较, 选出最优模型.

(1) 模型构建.

模型 4. 利用其中 1 组学生的试卷作答反应进行探索性因素分析, 以构建模型 4. 采用主成分分析法对因子进行抽取, 转轴法采用正交旋转法得到 3 个因子, 解释率总变异量为 49.7%, 第 1、2 和 3 个

因子上分别有20、14和12道题,每个因子上归属题目数量符合要求.因子1方差贡献率为20.3%,所包含的20道题中,有19道题主要涉及对物理科学知识、概念及原理和规律的识记.故该因子命名为科学知识与方法.因子2方差贡献率为16.1%,所包含的14道题难度系数均 >0.900 .故该因子命名为难度.因子3方差贡献率为13.3%,所包含的12道题中,有9道题涉及将物理科学知识及原理运用到题目情境中以解决问题,这些题目需要学生在基本公式及概念的基础上,结合题目给出的情境,对题目进行深入分析,运用科学思维予以解答.故该因子命名为科学应用.

模型5~7.经过因子结构的探索性因素分析,试卷内包含难度因子,则本研究在模型2的基础上,增加难度因子,划分为物质、运动和相互作用、能量和难度的4个因子模型(模型5).在模型3的基础上,增加难度因子,划分为了解、理解和难度的3个因子模型(模型6).此外,考虑到在模型4中难度因子所包含题目涉及到概念的基本运用,并且,科学知识与方法及难度2个因子之间存在正相关($r=0.846\sim 0.970, P<0.05$).由此,将模型4中的难度因子与科学知识与方法因子合并,构建科学知识与方法及科学应用的2个因子模型(模型7).

(2)模型比较.

模型4与5~7比较.通过对模型4~7的验证性因素结果指标分析,4个模型的CFI、TLI均 >0.900 ,RMSEA均 <0.050 ,WRMR >1.000 .由此,模型7优于模型5和6,模型4与7指标接近.

模型4与1~3比较.综合分析每个模型的拟合指标,对构建的模型进行比较.与模型1~3相比,模型4的CFI(0.991)和TLI(0.991)均增大;RMSEA(0.024)和WRMR(1.029)均降低.由此,模型4优于模型1~3.

模型4与7比较.与模型7相比,模型4的CFI(0.991)和TLI(0.991)均增大;RMSEA(0.024)和WRMR(1.029)均降低.进一步, χ^2 检验和模型拟合指数显示,2个模型之间差异显著($\Delta\chi^2/\Delta df=51.7, P<0.01$).结合2个模型的各项评价指标比较结果,模型4的观测数据与模型拟合更好($\chi^2=1\,610.6, df=986, P<0.05$;RMSEA=0.024, CFI=0.991, TLI=0.991, WRMR=1.029).由此,选取模型4为本研究的最佳模型.

3 结论与启示

研究显示,试卷的结构基本上是从科学知识与方法、科学应用2个因子划分.这与当前的课标要求是一致的.

物理学科作为科学教育的重要组成部分,其学科教育的核心在于培养学生的科学素养.通过对比国际大型测评PISA、NAEP和TIMSS项目科学测评框架,各项目将重点均放在了科学能力上.尽管测试指标不尽相同,但都包含了科学内容、科学方法2个重要方面.课标中所要求的义务教育阶段物理学科评价的框架包含了科学内容、认知和情感态度价值观.本研究从试卷作答的结果入手进行分析,明确了当前中考物理试卷考察的重点在科学知识与方法、科学应用2个方面.中考物理试卷从整体上可以体现对物理核心概念的认识和了解,及将物理科学应用、物理思维迁移运用到身边的生活及社会实践中.这表明当前物理学业水平考试评价重视能力立意,旨在考察学生真实的科学学科能力.基于此,本研究提出以下建议:

(1)明确测试框架.

在研究过程中,部分试卷题目在多个因子上均有载荷,且载荷相当.这与测试框架不无关系.如果能够有细化、可操作化的测试框架,则可以提高题目命制的精准度.当前,我国中考物理命题实行区域划分,各地区以课标为主要依据自行命题.由于课标属于指导性、上位的政策文件,无法对知识、能力等各维度进行细致划分,对各测查维度没有良好的结构化及操作化描述,为了更好地落实课标中对知识、能力的要求,在中考物理命题之初,需要围绕物理核心素养,依据课标设计出具有可操作性的指标框架,并以此指导命题.

(2)调整评价指标.

科学知识与方法、科学应用可能是试卷中潜在的结构,与当前已有内容及认知维度划分符合程度不高.因此,可以尝试借鉴国际已有评价项目对科学素养概念的理解,结合我国科学教育及物理教育发展的现实,以界定科学核心素养、明确物理学科核心素养为研究起点.并在此基础上,开展文献研究、现场调研、收集物理教育、科学教育、心理学、教育学、教育统计与测量学等方向广泛专家的意见.此外,通过多轮小规模测试,

获得进一步调整并细化评价指标的实证依据. 通过系统的评价指标研制, 力求指标明细化, 深入设计形成以培养物理核心素养为目标的评价指标体系.

(3) 提高试题命制技术.

试卷中难度水平在 0.900 以上的题目成为了单独的因子出现. 进一步分析这些题目, 学生基于个

体成长经验或生活常识即可正确作答, 与学科拟调查的内容和水平要求关系较弱. 由此, 在未来试题设计开发时, 从指标框架的制定、细目表的制定到命题人员的学科背景、统计测量背景 and 心理学背景的要求等多方面考虑, 注重拟考察指标的内涵和结构分析, 更好地使题目反映学科能力指标的本质属性, 进一步提高试卷质量.

参 考 文 献

- [1] 孙可平, 邓小丽. 理科教育展望[M]. 上海: 华东师范大学出版社, 2002: 272 - 273.
- [2] 胡献忠. 新版英国《国家科学教育课程标准》及其启示[J]. 全球教育展望, 2001(3): 44 - 49.
- [3] 朱汝葵. 着眼于提高学生科学素养的初中化学教材分析[D]. 上海: 华东师范大学, 2009.
- [4] Achieve. Next generation science standard[EB/OL]. (2017-04-11) [2018-11-10]. <http://www.nextgenscience.org/>.
- [5] National Commission on Excellence in Education. A nation at risk: the imperative for educational reform[J]. The Elementary School Journal, 1983, 84(2): 112 - 130.
- [6] SALGANIK L H, DOGAN E. National assessment of educational progress[M]. Netherlands: Springer, 2014.
- [7] 张俊杰. 新疆物理中考试卷与课程标准(实验版)的一致性研究[D]. 贵阳: 贵州师范大学, 2015.
- [8] 王树红. 物理中考试卷与课程标准的一致性研究[D]. 武汉: 华中师范大学, 2015.
- [9] 党越. 新课标下 2013—2015 年贵阳市中考物理试卷分析[D]. 贵阳: 贵州师范大学, 2016.
- [10] 袁令民. 中考物理科学探究能力评价特点研究[D]. 重庆: 西南大学, 2010.
- [11] 王秀梅. 天津市近十年中考物理试题的研究[D]. 天津: 天津师范大学, 2015.
- [12] 孔建斌, 武昌, 荆美婵. 实验探究求创新 强化过程见成效: 2012 年山西省中考物理命题思路解读[J]. 教育理论与实践, 2012, 32(29): 7 - 10.
- [13] 郭金, 杨净. 对 2013 年河北省物理学科中考命题的分析及复习建议[J]. 教育实践与研究(B), 2014(9): 54 - 57.
- [14] MILLER J D. Scientific literacy: a conceptual and empirical review[J]. Scientific Literacy, 1983, 112(2): 29 - 48.
- [15] Organisation for Economic Cooperation and Development. The PISA 2003 assessment framework: mathematics, reading, science and problem solving knowledge and skills[M]. Paris: [s. n.], 2003: 135 - 139.
- [16] 王文静, 王存宽, 周勇. PISA 2015 科学素养内涵变化、价值取向及启示[J]. 宁波大学学报(教育科学版), 2015, 37(3): 120 - 124.
- [17] 刘克文, 李川. PISA 2015 科学素养测试内容及特点[J]. 比较教育研究, 2015(7): 98 - 106.
- [18] MULLIS I V S, MARTIN M O, RUDDOCK G J, et al. TIMSS 2011 assessment frameworks[M]. Amsterdam: International Association for the Evaluation of Educational Achievement, 2009.
- [19] National Assessment Governing Board. Science framework for the 2011 national assessment of educational progress[M]. Washington: National Assessment Governing Board, 2010.
- [20] 中华人民共和国教育部. 义务教育物理课程标准(2011 年版)[S]. 北京: 北京师范大学出版社, 2011.
- [21] 续佩君. 中学物理教学中学生能力的量化研究(2): 中学生的物理能力教育目标[J]. 首都师范大学学报(自然科学版), 2000, 21(3): 21 - 24.
- [22] 肖骁, 邢红军. 高中《物理课程标准》中的科学方法显化研究[J]. 首都师范大学学报(自然科学版), 2011, 32(4): 29 - 34.

Revelation and Analysis of the Factor Structure of the Physics Test Papers for Junior High School Entrance Examination

ZHAI Yanfang¹ XING Hongjun²

(1. China Basic Education Quality Monitoring and Cooperative Innovation Center, Beijing Normal University, Beijing 100875;

2. College of Teacher Education, Capital Normal University, Beijing 100048)

Abstract: Based on the compulsory education physics curriculum standard and examination outline, through the factor analysis of the real data of the physics papers for junior high school admissions examination of Beijing, the three-factor structure of the test paper was discovered, and different models were further constructed. According to comparison and theoretical analysis, the optimal model was found. Thus, the real structure of the test paper was revealed, and it provided an available revelation for developing the physics test for junior high school entrance examination.

Keywords: junior high school physics; factor analysis; junior high school entrance examination

(上接第52页)

Research on Negative Image of Scenic Spots and Related Strategies Based on Network Comments ——a Case of Dalian Sunasia Ocean World

SONG Shunan CHANG Gaixin

(College of History Culture and Tourism, Liaoning Normal University, Dalian Liaoning 116000)

Abstract: With the rapid development of the internet, the relationship between tourism and the internet gets closer, therefore online comments become important references of tourism research. Online comments, which reflect tourists' impression and view, are important reflections of the image of the scenic area. However, the negative comments, which reflect the adverse perceptive image of the scenic area, are direct instructions for the scenic area to find problems and improve management. This paper uses the content analysis to study the negative comments on Dalian Sunasia Ocean World on ctrip's website. The comments are structured to form four first-level categories of overall experience, core attraction, specific venue and projects, operation and management, and nineteen second-level categories. According to the number and main points of negative comments of each category, the negative image of Sunasia Ocean World perceived by tourists and the reasons for its formation are summarized. Meanwhile, suggestions are put forward such as highlighting the characteristics, improving performance quality, strengthening venue maintenance and improving management. By exploring the main perceptual factors and forming reasons of the negative image of the ocean performance scenic area, it enriches the research content of the scenic area image, and provides a theoretical basis for the image enhancement of such scenic areas.

Keywords: online comments; negative image of scenic spot; Dalian Sunasia Ocean World; content analysis