

基于虚拟设计性物理实验的科学探究能力评价研究

——以美国科罗拉多州立大学虚拟设计性物理实验为例

陈雅馨 张志华 方 恺

(同济大学物理科学与工程学院, 上海 200092)

摘 要 科学探究能力是体现科学素养水平的重要指标,科学探究能力的培养已成为当前国际科学教育改革核心理念。物理课程标准对科学探究及大学物理实验能力作出了明确的要求。但是,目前在“设计性物理实验”中表现出的科学探究能力的评价体系并不完整。为此,本研究构建了基于虚拟设计性物理实验的科学探究能力评价量表及结构方程模型,该模型在有效性和合理性等方面均达到标准。同时本研究根据模型拟合结果对科学探究能力各维度设置了相应权重。基于此模型进行实证研究,结果表明,不同性别的学生在科学探究能力方面存在显著差异,而不同年级的学生则没有显著差异。此外,学生已有的知识水平对科学探究能力下“分析与论证”维度起到了较为显著的调节作用。

关键词 科学探究能力评价;虚拟设计性物理实验;结构方程模型;调节效应

RESEARCH ON THE EVALUATION OF SCIENTIFIC INQUIRY ABILITY BASED ON VIRTUAL DESIGN PHYSICS EXPERIMENTS

—TAKING THE VIRTUAL DESIGN PHYSICS EXPERIMENT OF
COLORADO STATE UNIVERSITY AS AN EXAMPLE

CHEN Yaxin ZHANG Zhihua FANG Kai

(School of Physics Science and Engineering, Tongji University, Shanghai 200092)

Abstract Scientific inquiry skills are an important indicator of scientific literacy, and the cultivation of such skills has become a core concept in current international science education reform efforts. Physics curriculum standards explicitly require the development of scientific inquiry and university-level physics laboratory skills. However, the current evaluation system for scientific inquiry ability demonstrated in “design-based physics experiments” is not complete. Therefore, this study constructed a scientific inquiry ability evaluation scale and a structural equation model based on virtual design-based physics experiments. The model achieved standards in validity and rationality. At the same time, corresponding weights were assigned to each dimension of scientific inquiry ability based on the model fitting results. Empirical research based on this model indicated that there were significant differences in scientific inquiry ability between students of different genders, but no significant differences were found among students of different grades. In addition, the existing knowledge level of students had a significant moderating effect on the “analysis

收稿日期: 2023-06-02

基金项目: 教育部产学研协同育人项目(202002123019), 同济大学实验教改专项基金项目(2020-51, 2021-49)。

通讯作者: 张志华, 同济大学物理科学与工程学院副教授, zzhtj@tongji.edu.cn。

引文格式: 陈雅馨, 张志华, 方恺. 基于虚拟设计性物理实验的科学探究能力评价研究——以美国科罗拉多州立大学虚拟设计性物理实验为例[J]. 物理与工程, 2024, 34(1): 120-129.

Cite this article: CHEN Y X, ZHANG Z H, FANG K. Research on the evaluation of scientific inquiry ability based on virtual design physics experiments—Taking the virtual design physics experiment of Colorado State University as an example[J]. Physics and Engineering, 2024, 34(1): 120-129. (in Chinese)

and reasoning” dimension of scientific inquiry ability.

Key words scientific inquiry ability evaluation; virtual design physics experiments; structural equation model; moderating effect

随着 21 世纪科学技术的迅速发展,科学素养已经成为人才必备的核心素养之一。科学探究能力的高低是科学素养水平的直接体现,许多国家和教育组织已将科学探究能力纳入他们的课程和评估之中,现今的科学教育已经从单纯关注事实和概念记忆转变为注重培养学生的高阶思维技能。在实验教学过程中,科学素养能在实验的每一个步骤中得到很好的培养和提升^[1],作为一门大学理工科专业的公共基础课程,物理实验课程教学重视学生科学探究能力培养,“设计性物理实验”是物理课程中重要的内容。

设计性物理实验衔接了基础教学实验与高阶科学研究实验,其教学以学生为主体,由学生自己动手设计,在实验过程中,思考和认识实验现象和实验结果所反映的物理本质,从而形成物理概念,揭示物理规律。相较于教学内容较为固定的基础实验,设计性物理实验的教学目标重在培养学生科学探究能力、开放性思维和创新意识^[2-3]。课程标准对科学探究及大学物理实验能力作出了明确的要求,其中在“设计性物理实验”中表现出的科学探究能力的评价体系需要进一步明确和研究。因此,建立一个基于设计性物理实验的科学探究评价模型是当前大学物理教育研究中的一个重要课题。

本研究利用结构方程模型构建了基于虚拟设计性物理实验的科学探究评价体系,模型在有效性、合理性等方面均达到标准。同时,本研究利用此模型进行了实证研究,探索当代大学生科学探究能力现状,并通过实验结论得出改进科学探究教学和评价的建议以及对该评价模型的优化方向。

1 文献综述

1.1 设计性物理实验

设计性实验中,实验方法的可拓展性和实验结果的不确定性要求学生具备探究的意识以及掌握探究的基本原则和技巧,设计性实验有助于激发学生的探索热情和提高科学探究问题的能力^[4]。

根据设计性物理实验类型属性进行层次划分,大致可分为验证型、研究型、应用型和高新

型^[2]。验证型是指学生独立设计出一套新颖且有效的方法来验证某一物理现象或物理定律。研究型强调探索物理量之间的关系、物理学的内在联系以及对物理实验进行改革、拓展和创新的研究。应用型重在体现其设计的应用性,学生需要设计出某一产品或某一测量方法能为实际所用。高新型主要鼓励和选拔优秀学生参与到物理学中某些热点、前沿领域的研究和开发活动。

根据设计性物理实验的操作类型可以分为虚拟实验和实做实验,相关研究表明^[5],在正确的环境中使用虚拟实验室作为教育工具,无论是培养学生使用物理仪器方面,还是帮助学生理解知识概念方面比实做实验室更有效。

因此,本研究采用美国科罗拉多州立大学开发的电磁学虚拟实验室^[8](如图 1 所示)作为设计性物理实验实证研究载体。国际上基于此实验室已经进行了多项研究,Keller, C. J. 等人证实该虚拟实验室在大学本科物理教育中有效^[6]; Katherine Perkins 等人利用访谈法进行实证研究,发现 70% 的学生认为该虚拟实验室能帮助他们理解复杂的电磁学物理知识^[7]。虚拟实验室具有多种电路元件及配套测量工具,在该线上实验室中可以完成的设计性物理实验类型主要为:验证型和研究型,学生可自主选择相关电路元件验证某一物理现象和物理定律或者研究物理量之间的关系。



图 1 美国科罗拉多州立大学虚拟实验室^[8]

1.2 科学探究能力评价体系

科学教育的教育理念起源于国外,国际典型的科学探究能力评价的方法主要有国家层面的评价项目及三大国际著名的学业评级项目。国家层面上,主要有英国的学业评价机构(APU)、普通教育证书(GCE)以及美国的 PADI 科学探究能力评

价模式。三大国际著名学业评级项目为经合组织 OECD 组织实施的国际学生评价项目 PISA、1995 年由国际教育成就评价协会发起的国际数学和科学评测趋势项目 TIMSS 以及 1969 年由美国国家教育评测中心开始实行的全国性评测 NAEP。

APU 认为,科学探究是一种以解决问题为目标的过程,它通过获取有效证据和反思评价来不断调整和改进整个科学探究过程,APU 将科学探究能力要素分为 6 个方面^[9]:符号表征的使用、器材和测量工具的使用、观察能力、解释和应用、探究设计、整体探究表现。PISA2015 提出了科学素养的概念,它包括科学知识和科学技术,而科学能力则是其中的核心。科学能力主要分为三种^[10]:第一种是能够解释各种科学现象,对其进行认识和评价;第二种是能够评价和设计科学探究,即提出科学问题并找到解决方法;第三种是能够科学地解释数据和证据,分析不同的科学数据和参数表达方式,并得出准确的结论。

马永双采用 PTA 量表法,通过对学生科学探究能力活动及能力进行要素分析,并借鉴多位高校教师的建议和相关文献,创建了科学探究能力的 PTA 评价量表,使得科学探究能力的评价从抽象化变得具体化^[11]。该评价量表在多篇硕士学位论文中得到了应用,例如郭静的《中学化学教师及化学师范生科学探究能力发展现状的调查研究》^[12]。

从以上国内外所提到的评价体系来看,都涉及了科学探究领域的实践内容,比如 APU 将科学探究能力要素分为 6 个方面^[9];PISA2015 提出了 3 种能力要素^[10]。由具体的评价内容来看,以上的评价体系都涉及了以下的 6 个维度:

(1) 提出问题:对给定情境中的现象进行观察,从而提出可以研究的科学问题;

(2) 猜想与假设:针对自己提出的问题,作出合理的假设;

(3) 实验设计:在情境中设计实验方案;

(4) 数据收集:对实验数据的收集和处理;

(5) 分析与论证:根据收集到的数据对假设进行验证,并且下结论;

(6) 评价与反思:反思自己的实验过程以及需要改进的地方。

所以本研究将根据以上总结出的几个维度进行基于设计性物理实验的科学探究能力量表设计。

1.3 科学探究能力评价形式

国际上评价学生科学探究能力主要有四种形

式:观察法、纸笔测验、工作单和计算机模拟^[13]。观察法是老师们在一旁观察学生的探究活动,并根据观察结果给出定性和定量的评价。虽然这是被视为最有效的评价方式,但成本较高^[13],纸笔测验是以测试的形式考察学生的探究能力,虽然成本较低,但题目之间关联度不高,与真实情境表现出的探究能力有差异^[14]。工作单则是让学生实际进行探究并将结果写在一张单子上,教师进行延时评分,其成本较低且操作方便,因此得到广泛地应用^[14]。计算机模拟是一种新的评价方式,学生在计算机模拟的探究情境中进行操作并接受评分,具有成本低、评价可观性强和及时保存等优点,不过,其前期开发成本较高,且需要被测试者具备一定的计算机操作能力^[15]。

本研究采用工作单的评价形式,根据开放程度的多少,可以将工作单分为三种:引导性、开放性和结构性工作单,国际上也有一些著名的教育评价专家对工作单的效度进行了实证研究,发现其效度是依次升高的,均达到了 0.67^[16]。本研究采用效度最高的结构性工作单,将工作单开发成电子工作单,嵌入物理实验平台,学生进行设计性实验时填写工作单,研究者按照构建的科学探究能力评价量表进行延时评分。

2 研究设计

2.1 研究对象

本研究以华东地区 T 大学 203 名学生作为研究对象。通过收集这 203 名学生在设计性物理实验中填写的知识水平前测问卷及结构性工作单数据,并进行整理,清除无效数据后,最终筛选出 189 名学生作为研究对象。其中,大一学生为 80 人,大二学生为 109 人;男性 134 人,女性 55 人。

2.2 科学探究能力评价量表设计

科学探究能力必须外显为具体的、可测量的行为,才能进行有效地评价。科学探究能力是隐变量,不能直接对其评价,要将这些要素外显为学生的行为特征,才能完成评价^[17]。将科学探究能力划分为以下 6 个二级维度:提出问题、猜想与假设、实验设计、数据收集、分析与论证、评价与反思,每个二级维度需要考虑其 2~3 个具体的表现水平。这些水平应基于以下几个因素:系统性、深刻性、相关性和创造性,其表现水平分别对应“孤立—联系—系统”“局部—完整”“表象—本质”“无

关—相关”“简单情境—复杂情境”等^[18]。

基于设计性物理实验的科学探究能力评价体系中,最核心的二级维度为:实验设计,该维度能反映学生在进行设计性物理实验时真实的实验设计能力;为了对学生的学习成果进行分类,本研究采用 SOLO (Structure of the observed learning

outcome) 分类理论,将学生的可观察学习成果结构分为五个层次:前结构、单点结构、多点结构、关联结构和抽象扩展结构^[19]。每个层次给出不同的等级分数,这样学生的回答质量就可以被量化,分数也可以作为评价的依据^[19]。具体的评价维度和其对应的外显水平得分如表 1 所示。

表 1 基于设计性物理实验的科学探究能力评价量表

二级维度	三级维度	具体表现(外显)	得分
提出问题(A)	提出问题个数、与情境相关性(A1)	提出三个及三个以上与给定情境相关的问题	3
		提出两个与给定情景相关的问题	2
		提出一个与给定情景相关的问题;或提出两个或两个以上的问题,但与情景相关性小	1
		提出一个或未提出问题,且与情境相关性小	0
	问题的科学性(A2)	问题表述清楚、明了,符合科学规范	2
		问题表述清楚、但不符合科学规范	1
		提出问题不符合科学规范	0
	问题的逻辑性(A3)	提出问题表述、逻辑清晰	2
		提出问题表述基本清晰,部分逻辑混乱	1
		提出问题表述混乱,逻辑不清晰	0
猜想与假设(B)	假设的个数、与问题的相关性(B1)	提出两个假设以上,且与问题相关	2
		提出一个假设,且与问题相关	1
		提出的假设无论个数都与问题不相关	0
	猜想与假设的合理性(B2)	猜想与假设合理、能在实验中进行探究	2
		猜想与假设部分合理,部分能在实验中探究	1
		猜想与假设不合理、无法在实验中进行探究	0
	猜想与假设的质量(B3)	猜想与假设表述清晰,能给出合理的理由	2
		猜想与假设表述较为清晰,能给出理由但不充分或不能给出理由	1
		猜想与假设表述不清晰,不能给出合理的理由	0
实验设计(C)	识别自变量和因变量(C1)	能识别自变量和因变量,语言描述具体、精确、变量有名称和符号表示	3
		能识别自变量和因变量,语言描述较为具体,精确、变量有名称表示	2
		能识别自变量和因变量,语言描述含糊	1
		无法识别自变量和因变量	0
	识别控制变量(C2)	能识别多类控制变量,语言描述具体、精确、变量有名称和符号表示	3
		能识别多类控制变量,语言描述较为具体,精确、变量有名称表示	2
		只能识别 1 类控制变量或识别多类变量但语言描述含糊	1
		无法识别控制变量	0

续表

二级维度	三级维度	具体表现(外显)	得分
实验设计(C)	对变量下操作性定义(C3)	能对自变量和控制变量下操作性定义,语言描述具体、精确	3
		能对自变量和控制变量下操作性定义,语言描述较为具体、精确	2
		只能对自变量下操作性定义,语言描述模糊	1
		不能对变量下操作性定义	0
	变量重复测量方法及数据获取时间(C4)	知道对因变量和控制变量重复测量的(方法)或知道何时记录数据,表述清晰,有具体的时间点和次数说明	3
		知道对因变量和控制变量重复测量的(方法)或知道何时记录数据	2
		知道对因变量重复测量的(方法)或知道何时记录数据	1
		没有作答或无相关内容	0
	实验设计思路(C5)	答出完整且顺序合理的实验步骤,且有对具体注意事项的描述	3
		能够答出完整且顺序合理的实验步骤,但缺少对具体注意事项的描述	2
		答出部分顺序合理的实验步骤	1
		没有任何信息或只有无效的操作步骤	0
数据收集(D)	数据的完整性(D1)	收集的数据完整	2
		收集的数据部分完整	1
		收集的数据完全缺失	0
	数据的利用(D2)	利用数据分析问题,并对数据进行处理和计算	2
		利用所有数据,但没有对数据进行处理和计算	1
		数据利用不充分	0
分析与论证(E)	是否利用数据分析工具(D3)	利用数据分析软件或 excel 图等进行分析	1
		没有利用数据分析软件或 excel 图等进行分析	0
	结论的科学性(E1)	结论符合科学结果	1
		结论偏离科学结果	0
	结论的逻辑性(E2)	能用完整的物理语言表述结论,结论逻辑清晰	2
		能用较为完整的物理语言表述结论,结论逻辑较为清晰	1
		不能用完整物理语言表述结论	0
评价与反思(F)	评价与反思(F)	能发现实验中的不足,并给出改进方法	2
		能发现实验中的不足,但不能给出改进方法	1
		无法发现实验中的不足	0

2.3 研究平台搭建

本研究采用 html5、JavaScript、Springboot、MySQL 等编程技术,开发了 PC 端物理实验平台,学生可以在该平台上完成知识水平的前测及设计性物理实验的结构性工作单^[20](如图 2、3 所示)。平台的搭建提升了数据收集的效率,可以满足上百名学生同时进行设计性物理实验。

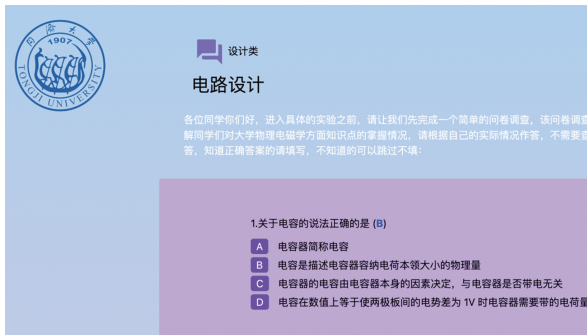


图2 知识水平前测页面截图

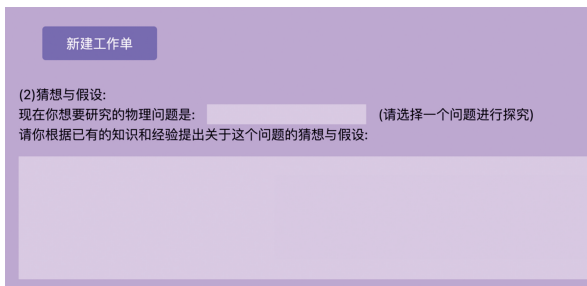


图3 部分结构性工作单截图^[20]

3 研究结果

3.1 模型检验

本研究共收集了 189 份学生设计性物理实验工作单数据,用拟定的科学探究能力评价量表进行评分,利用 SPSS 软件与 AMOS 软件分别对基于虚拟设计性物理实验的科学探究能力测量模型的信效度以及其结构模型的拟合度进行检验。为了进一步确定指标权重,确保模型能够运用到教学实践中,利用各变量间的标准化系数来确立潜变量和观测变量的权重。

1) 测量模型信效度检验

使用 SPSS 27.0 统计软件分析初始测量模型所得 Cronbach's α 系数 = 0.839 > 0.8,说明整体量表的可靠性较好。在 Bartlett 球形检验中, KMO(Kaiser-Meyer-Olkin, 检验统计量)系数 = 0.782 > 0.7 且 $p = 0.000 < 0.05$,符合球形检验标

准。结合 KMO 值和 Bartlett 球形检验值两项指标综合评估,本量表信效度良好,适合进一步分析。

2) 结构模型拟合度检验

用 AMOS 软件对假设模型和各个因子之间的关系进行检验,根据检验结果对假设模型进行多次修订。运行结果如表 2 所示,绝对适配指数、增值适配指数以及简约适配指数均在拟合指数评价标准范围之内,这说明结构模型的拟合度较好,可以进行结果解释。

表2 结构模型拟合指数

类型	指数	拟合结果	拟合指数评价标准
绝对适配指数	CMIN/DF	2.016	<3 较优, <5 可接受
	RMSEA	0.074	<0.08
增值适配指数	CFI	0.915	>0.9
	IFI	0.917	>0.9
简约适配指数	PNFI	0.698	≥ 0.5
	PCFI	0.754	≥ 0.5

3) 权重设置

图 4 是完整的结构方程模型,其中观测变量、

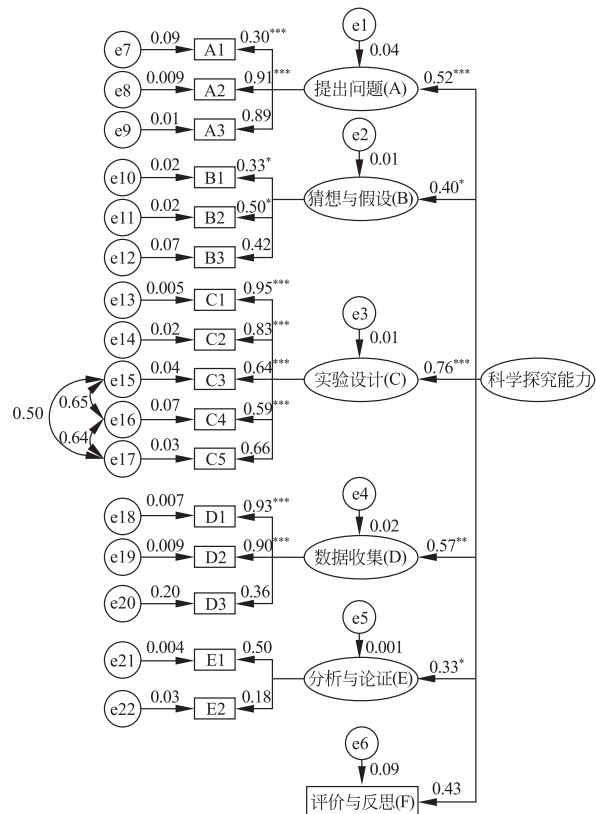


图4 基于设计性物理实验的科学探究能力模型

注: ***表示 $P < 0.001$; **表示 $P < 0.01$; *表示 $P < 0.05$

一阶潜变量、二阶潜变量在表 1 中已做定义,各变量之间用单箭头直线连接,标准化路径系数的数值越大则说明变量之间的影响效果越大,各个路径之间的 P 值均显著,即一阶潜变量与二阶潜变量之间,二阶潜变量与观测变量之间均通过了检验(根据结构方程相关理论,一组观测变量中至少一个观测变量需预设其路径系数为 1,预设路径系数为 1 的变量,无需讨论其显著性^[21]) $e(e1 \sim e22)$ 为干扰变量,它与观测变量之间的数字代表测量误差。

为进一步确定指标权重,确保模型能够将学生的科学探究能力量化,需要对各类评价指标的标准化系数按照式(1)进行归一化处理^[22]。 W 为模型指标权重, c 为指标标准化系数。

$$W_i = \frac{c_i}{\sum_i c} \tag{1}$$

表 3 是完整的科学探究能力评价模型,将 6 项二级维度的评价模型归一化处理后可以得到一阶维度“科学探究能力”的评价模型为:科学探究能力 $=0.17 \times$ 提出问题(A) $+0.13 \times$ 猜想与假设(B) $+0.25 \times$ 实验设计(C) $+0.19 \times$ 数据收集(D) $+0.12 \times$ 分析与论证(E) $+0.14 \times$ 评价与反思(F)。

利用“科学探究能力”评价模型对 189 名同学的科学探究能力进行评分,因数据已归一化处理,评分范围在 0~1 分之间,并应用于后续数据分析。

3.2 实验数据分析

1) 性别、年级因素对科学探究能力影响

利用 SPSS 27.0 软件,将男、女两组同学的科学探究能力水平及其 6 个二级维度进行独立样本 T 检验,结果如表 4 所示。1 组为男性,2 组为女性。发现男性、女性在设计性物理实验中表现出的科学探究能力有显著性差异($P=0.014<0.05$)、且具体差异表现在二级维度中的:实验设计(C)维度($P=0.010<0.05$),女生在实验设计维度均分高于男生,其余二级维度中均没有显著性差异。

利用 SPSS 27.0 软件,将大一、大二两组同学的科学探究能力水平进行独立样本 T 检验,结果如表 5 所示。1 组为大学一年级本科生,2 组为二年级本科生。发现两组学生在设计性物理实验中表现出的科学探究能力并没有显著性差异($P=0.793>0.05$)。可见,不同年级的学生所具备的科学探究能力并没有很大差别,即知识水平的增长并不一定能促进学生科学探究能力的提高。

表 3 基于设计性物理实验的科学探究能力评价模型

一级维度	二级维度	权重值	三级维度	权重值
科学探究能力	提出问题(A)	0.17	提出问题个数、与情境相关性(A1)	0.15
			问题的科学性(A2)	0.43
			问题的逻辑性(A3)	0.42
	猜想与假设(B)	0.13	假设的个数、与问题的相关性(B1)	0.26
			猜想与假设的合理性(B2)	0.40
			猜想与假设的质量(B3)	0.34
	实验设计(C)	0.25	识别自变量和因变量(C1)	0.26
			识别控制变量(C2)	0.23
			对变量下操作性定义(C3)	0.17
			变量重复测量方法及数据获取时间(C4)	0.16
	数据收集(D)	0.19	实验设计思路(C5)	0.18
			数据的完整性(D1)	0.42
			数据的利用(D2)	0.41
	分析与论证(E)	0.12	是否利用数据分析工具(D3)	0.17
			结论的科学性(E1)	0.74
	评价与反思(F)	0.14	结论的逻辑性(E2)	0.26
			评价与反思(F)	1.00

表 4 性别影响的科学探究能力及实验设计的 T 检验结果

维度	性别	平均值	<i>t</i>	自由度	显著性 (双尾)
科学探究能力	1	0.795	-2.477	146.036	0.014*
	2	0.840			
实验设计(C)	1	0.649	-2.631	125.675	0.010*
	2	0.729			

表 5 年级影响的科学探究能力的 T 检验结果

维度	年级	平均值	<i>t</i>	自由度	显著性 (双尾)
科学探究能力	1	0.811	0.263	186.848	0.793
	2	0.806			

2) 知识水平对“分析与论证”维度的调节效应

利用 Hayes (2012) 编制的 SPSS 宏中的 Model (Model 假设调节模型与本研究的理论模型一致), 自变量为分析与论证, 因变量为科学探究能力, 调节变量为学生已有的知识水平 (根据《大学物理》上册^[23]电磁学部分内容制定了知识水平的前测问卷, 共 6 道题目, 其中包括电磁学部分的 5 道客观题, 最后 1 题为自评题, 为了了解学生自己对电磁学知识点的认知评价, 学生的得分将作为学生已有知识水平维度), 对假设模型的调节效应进行检验, 结果如表 6 所示, 分析与论证及知识水平的乘积项对科学探究能力的预测作用显著 ($B = -0.38, t = -4.34, p < 0.001$), 说明学生的知识水平能够在分析与论证对科学探究能力的

预测中起调节作用。

进一步通过简单斜率分析 (如表 7、图 5 所示), 知识水平较低 ($M - 1SD$) 的被试, 分析与论证对科学探究能力具有显著的正向预测作用 ($simple\ slope = 4.20, t = 6.36, p < 0.001$); 而对于知识水平较高 ($M + 1SD$) 的被试, 分析与论证对科学探究能力产生负向预测作用, 且其预测作用较小 ($simple\ slope = -1.99, t = 3.47, p < 0.001$)。

分析出现高分组负向预测的原因为: 设计性物理实验本身是一个非常开放的实验, 学生可以自由设定探究题目。在已有知识水平较高的学生那组, 大部分高分组学生会选择更具有挑战性的探究题目 (例如: 探究为什么电容元件的交流电路中当电容和线路中电阻值较大时在开关闭合一段时间内电容器两端电压有效值非恒定? 低通滤波电路中传递函数与角频率关系验证等), 这些题目可能会超出他们已有的知识范围。因此, 这部分学生在其他维度表现优异, 分数较高, 但是在分析与论证部分可能会因为自己做出来的结果不佳, 导致分数较低。在高知识水平组中, 也会有部分学生选择非常简单的探究题目 (例如: 验证叠加定理、探究小灯泡工作时的内阻等), 他们对这些题目的结论已经非常了解和清晰, 因此他们在实验时并没有认真考虑其他维度, 其他维度得分较低, 但分析与论证维度得分较高, 这就导致了高知识水平组的斜率是负数的情况。该实验结果同时也引出了模型的优化方向: 可以将学生提出问题的难度系数融入模型中以更好地评价学生的科学探究能力。

表 6 模型回归方程检验

回归方程 ($N = 189$)		拟合指标			系数显著性	
因变量	预测变量	<i>R</i>	R^2	<i>F</i>	<i>B</i>	<i>t</i>
科学探究能力		0.39	0.15	16.45***		
	分析与论证				0.39	5.73***
	知识水平				0.04	0.53
科学探究能力		0.48	0.23	18.29***		
	分析与论证				0.65	7.36***
	知识水平				0.09	1.36
	分析与论证 $\times$ 知识水平				-0.38	-4.34***

注: *** 表示 $p < 0.001$ 。

表 7 高、低知识水平组调节效应检验

知识水平	Effect	标准误差	t	P	LLCI	ULCI
-0.21 (M-1SD)	1.74	0.27	6.36	<0.001	1.20	2.27
0 (M)	1.08	0.15	7.36	<0.001	0.79	1.37
0.21 (M+1SD)	0.42	0.12	3.47	<0.001	0.18	0.66

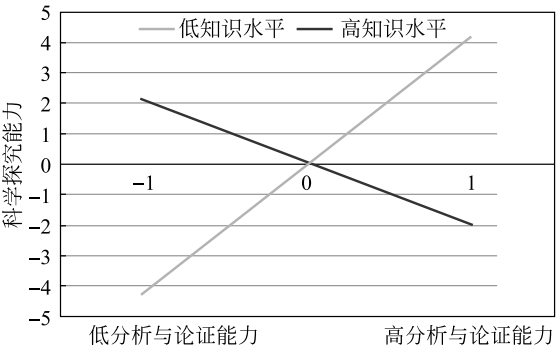


图 5 知识水平在分析与论证与科学探究能力关系中的调节作用

4 讨论与启示

由研究结果可以得出,不同性别的学生所具备的科学探究能力有显著差异,且具体差异表现在二级维度中的:实验设计(C)维度,女生在实验设计维度均分高于男生。顾乃景等人也曾在研究中发现,性别因素对学生的科学探究能力存在影响,且女生的科学探究能力均分高于男生^[24]。不同年级的学生所具备的科学探究能力没有显著差异,年级越高的同学并不一定科学探究能力越强。

研究结果还发现,学生已有的知识水平对科学探究能力下的“分析与论证”维度有显著的调节效应,知识水平较低的被试者,分析与论证对科学探究能力具有显著的正向预测作用,知识水平较高的被试者,分析与论证对科学探究能力产生负向预测作用,且其预测作用较小。这说明在低阶的科学探究能力活动中,被试者已有知识水平较低的情况下,纸笔检验或实验报告(两种形式主要考察学生分析与论证的能力)的检验形式,一定程度上可以反映学生的科学探究能力,但在高阶的科学探究能力活动中,被试者已有知识水平较高的情况下,会削弱分析与论证对科学探究能力的

预测作用。在此情境下,教师不能只依据学生的实验报告或纸笔检验来评价学生的科学探究能力,要结合多维度的评价体系,例如使用本研究所构建的科学探究能力评价模型进行评价。

本研究利用结构方程模型构建了基于设计性物理实验的科学探究能力评价模型,该模型在有效性、合理性等方面均达到标准,科学化程度较高,但其中不足之处在于,设计性物理实验作为一个完全开放性的实验,学生自己所设定的探究题目也应存在难易度区分。后续研究可以将学生的探究题目的难度系数融入该模型中以更好的评价学生的科学探究能力,而探究题目的难度系数如何定义则需进一步研究。

5 结语

本研究构建的科学探究能力评价模型以多维度的评价方式力求更为全面地反映学生的科学探究能力,教师可以通过评价结果及时发现学生面对设计性物理实验中的薄弱环节并加以引导,帮助学生开拓思维,挖掘创造潜力,进一步提升科学探究能力。在今后的大学物理设计性实验教学中,我们将进一步优化科学探究能力评价模型,同时研究评价模型对实做设计性物理实验的实用性,从而助力物理实验教学改革。

参 考 文 献

[1] 沈秀,张黎,戎红仁,等. 深化实验教学改革提升大学生科学素养[J]. 实验室研究与探索, 2012, 31(3): 163-165, 192.
SHEN X, ZHANG L, RONG H R, et al. Deepening the Reform of Experimental Teaching to Improve College Students' Scientific Literacy[J]. Laboratory Research and Exploration, 2012, 31(3): 163-165, 192. (in Chinese)

[2] 漆建军. 设计性物理实验的设计与实施研究[D]. 广东: 华南师范大学, 2004.

[3] 何深知,崔纯莹,段薇,等. 在设计性实验中培养学生的综合能力[J]. 实验室研究与探索, 2009, 28(6): 134-136.
HE S Z, CUI C Y, DUAN W, et al. Cultivating Students' Comprehensive Abilities in Design-Oriented Experiments [J]. Laboratory Research and Exploration, 2009, 28(6): 134-136. (in Chinese)

[4] 吴肖,熊建文. 物理设计性实验中的学生实验行为[J]. 实验室研究与探索, 2013, 32(1): 108-110.
WU X, XIONG J W. Students' Experimental Behaviors in Design-Oriented Physics Experiments[J]. Laboratory Research and Exploration, 2013, 32(1): 108-110. (in Chi-

- nese)
- [5] FINKELSTEIN N D, ADAMS W K, KELLER C J. When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment[J]. *Physical Review Special Topics: Physics Education Research*, 2005, 1(1): 1-2.
- [6] KELLER C J, FINKELATEIN N D, PERKINS K K, et al. Assessing the Effectiveness of a Computer Simulation in Introductory Undergraduate Environments[J]. *AIP Conference Proceedings*, 2007, 883(1): 121-124.
- [7] PERKINS K, ADAMS W, DUBASON M. PhET: Interactive Simulations for Teaching and Learning Physics[J]. *The Physics Teacher*, 2006, 44(1): 18-23.
- [8] 玻金斯, 威曼. 虚拟物理实验[Z/OL]. [2023-07-20]. <https://phet.colorado.edu/en/simulations/circuit-construction-kit-ac-virtual-lab>
- [9] 钟媚. 英国 APU 科学探究能力测评及其启示[J]. *外国中小学教育*, 2010, (11): 38-42.
- ZHONG M. Evaluation of Scientific Inquiry Skills in APU in UK and its Enlightenment[J]. *Foreign Primary and Secondary Education*, 2010, (11): 38-42. (in Chinese)
- [10] 王文静, 王存宽, 周勇. PISA2015 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.
- WANG W J, WANG C K, ZHOU Y. Changes in Connotation and Value Orientation of Scientific Literacy in PISA 2015 and its Implications[J]. *Journal of Ningbo University (Education Science Edition)*, 2015, 37(3): 120-124. (in Chinese)
- [11] 黄臻, 唐劲军. 量表法评价中学生化学科学探究能力[J]. *化学教育*, 2007, (4): 30-32.
- HUANG L, TANG J J. Evaluation of High School Students' Chemical Scientific Inquiry Skills by Scale Method[J]. *Chemistry Education*, 2007, (4): 30-32. (in Chinese)
- [12] 郭静. 中学化学教师及化学师范生科学探究能力发展现状的调查研究[D]. 西安: 陕西师范大学, 2011.
- [13] 罗国忠. 科学探究能力的评价及其效度比较[J]. *教育科学*, 2013, 29(1): 10-13.
- LUO G Z. Evaluation of Scientific Inquiry Skills and Its Validity Comparison[J]. *Education Science*, 2013, 29(1): 10-13. (in Chinese)
- [14] 罗国忠. 科学探究能力的评价方法[J]. *教育科学*, 2007, (3): 7-10.
- LUO G Z. Evaluation Methods for Scientific Inquiry Skills [J]. *Education Science*, 2007, (3): 7-10. (in Chinese)
- [15] 顾乃景. 大学生科学探究能力评价研究[D]. 南京: 东南大学, 2016.
- [16] SHAVELSON R J, BAXTER G P, GAO X. Sampling variability of performance assessments[J]. *Journal of Educational Measurement*, 1993, 30(3): 215-232.
- [17] 李博宁. 高中生科学探究能力评价标准体系的研究[D]. 长春: 东北师范大学, 2008.
- [18] 刘东方. 科学探究能力表现模型的建构及其在评价中的应用[J]. *课程. 教材. 教法*, 2018, 38(9): 122-127.
- LIU D F. Construction of Performance Model for Scientific Inquiry Skills and Its Application in Evaluation[J]. *Curriculum, Teaching Material and Method*, 2018, 38(9): 122-127. (in Chinese)
- [19] 比格斯, 科利斯. 学习质量评价: SOLO 分类理论[M]. 高凌飏, 张洪岩, 译. 北京: 人民教育出版社, 2011.
- [20] 陈雅馨. 设计性物理实验结构性工作单[Z/OL]. [2023-07-20]. <https://docs.qq.com/doc/DY1lWQVFrSUtnUR6>
- [21] BYRNE B M. Structural equation modeling with AMOS: Basic concepts, applications, and programming[M]. 3rd ed. London: Routledge, 2016. 45-46.
- [22] 翟苗, 张睿, 刘恒彪. 高校混合式教学形成性评价指标研究[J]. *现代教育技术*, 2020, 30(9): 35-41.
- ZHAI M, ZHANG R, LIU H B. Research on Formative Evaluation Indicators of Blended Learning in Higher Education[J]. *Modern Educational Technology*, 2020, 30(9): 35-41. (in Chinese)
- [23] 毛俊健, 顾牡. 大学物理学. 上册[M]. 2 版. 北京: 高等教育出版社, 2013. 210-211.
- [24] 顾乃景, 伏干. 基于 PTA 量表的大学生科学探究能力评价研究—以南京地区高校为例[J]. *考试研究*, 2021, 17(6): 66-84.
- GU N J, FU G. Research on the Evaluation of Scientific Inquiry Skills of College Students Based on the PTA Scale: A Case Study of Universities in Nanjing[J]. *Examination Research*, 2021, 17(6): 66-84. (in Chinese)