

# 初中数学试题难度模型构建与分析

何子笛, 庄再博, 黄蔓, 王培蓓, 叶丽霞\*, 贾随军

浙江外国语学院数学系, 浙江 杭州

收稿日期: 2023年8月17日; 录用日期: 2023年10月17日; 发布日期: 2023年10月26日

## 摘要

在鲍建生五因素难度模型的基础上, 定义难度因素水平并改进了数学试题难度模型。选取宁波市部分中考题为例, 编制试题难度影响因素调查问卷, 再采用探索性因子分析、多元线性回归等统计方法分析调查数据, 得出试题难度的8个核心影响因素, 结合概念性要素、程序性要素、情境性及层次性要素3个维度, 构建初中数学试题难度线性模型, 并对模型的有效性进行了验证。

## 关键词

试题难度, 影响因素, 难度模型

# Construction and Analysis of the Difficulty Model for Middle School Mathematics Exams

Zidi He, Zaibo Zhuang, Man Huang, Peibei Wang, Lixia Ye\*, Suijun Jia

Department of Mathematics, Zhejiang International Studies University, Hangzhou Zhejiang

Received: Aug. 17<sup>th</sup>, 2023; accepted: Oct. 17<sup>th</sup>, 2023; published: Oct. 26<sup>th</sup>, 2023

## Abstract

On the basis of Bao Jiansheng's five factor difficulty model, the difficulty factor level was defined and the mathematical test paper difficulty model was improved. Taking some high school entrance examination questions in Ningbo City as examples, a survey questionnaire on factors affecting difficulty in test papers was developed. Exploratory factor analysis, multiple linear regression, and other statistical methods were used to analyze the survey data, and eight core influencing factors of test paper difficulty were identified. Combining three dimensions, which are conceptual elements, procedural elements, situational elements, and hierarchical elements, a linear model of difficulty in middle school mathematics test papers was constructed, and the effectiveness of the model was verified.

\*通讯作者。

## Keywords

### Test Paper Difficulty, Influencing Factors, Difficulty Model

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

教育部办公厅《关于做好2022年中考命题工作的通知》中提到,从对各地中考命题跟踪评估的情况看,一些地方还存在超标命题、试卷难度不合理、记忆性试题比例偏高等问题。在双减背景下,我们需要减少学生做题数量,但仍需保证学生的数学学习兴趣,最大程度促进学生全面发展,那么如何科学设计难度合理的作业和试卷来切实做到“减负提质”就显得尤为重要。因此,如何科学评估数学试题难度面临着极大的挑战。

有关数学学习任务的难度研究,国内外已有许多学者利用不同的难度模型来进行定性或定量的分析。1996年,邵志芳等人提出影响一般试题或问题难度的影响因素,包括未知参数的数量、中间未知参数不予提示程度、供选择的操作数等[1]。徐建乐认为试题的抽象水平、试题的综合程度和试题条件的隐蔽性是影响数学试题难度的三个主要因素[2]。2008年,杜明荣提出在具体情况下,试题的表达方式、条件的非充分性、问题目标的开放性是试题难度不可忽视的影响因素[3]。2014年,鲍建生等教授以“背景”“认识”“运算”“推理”“知识含量”五个难度因素结合已有研究建立了评估数学习题难度的模型,同时对每一因素进行水平划分,并通过该模型对中、美、法等国六套高中数学教材的例题难度进行比较分析[4]。2018年,武小鹏等人在鲍建生难度模型基础上进行了改编,将“思维方向”“是否含参”2个维度加入难度因素范畴进行编码考察[5]。2020年,李保臻等认为高考数学大题中通常涉及的多个小问之间的递进关系会对试卷综合难度产生影响,因此在武小鹏难度模型的基础上加入“梯度”因子,具体分“问题互不干扰”和“问题之间存在关联”两个层次[6]。

综上所述,上述模型大多仅能在教材比较的背景下使用,且鲍建生的难度模型仅能衡量题目的相对难度,只有在比较的情境下才有意义。同时,五个要素之间是否彼此独立或不相关,这点并未进行验证。所以建立绝对难度模型进行定量分析对于试题难度的研究具有重要的意义。

基于我国初中数学试题的实际情况,本文将诸多难度因素进行筛选概括,初步得到影响初中数学试题难度的影响因素——知识点数量、情境、运算、推理、认知、开放性、隐蔽性、梯度、未知参数数量共9个。此外,通过对问卷调查的结果进行因子分析,根据其相关性得出了三个主要成分因素。同时对各影响因素与难度进行一元线性回归分析,计算各影响因素相对于难度的系数,从而建立能够衡量任一数学试题绝对难度的有效模型,为试题命制、试题筛选和课堂教学提供可靠依据。该模型与鲍建生难度模型相比,具有更广泛的使用范围和普适性。

## 2. 研究工具

### 2.1. 研究框架

将以上内容作为理论基础,分四步构建研究框架,详见图1。第一步,采取文献分析法、对比分析等方法,查阅鲍建生教授建立的五因素模型以及后续模型发展等文献,确定研究方向为如何建立关于初

中数学试题难度的新模型。第二步，在鲍建生等教授五因素模型的基础上，结合专家意见，重新定义难度因素水平。并从宁波市历年中考卷中挑选出部分有代表性的题目编制调查问卷，向数学师范专业学生发放。第三步，剔除有明显数据缺失等情况的问卷后，对问卷进行信度检验，并对通过信度检验的问卷进行探索性因子分析，从实际数据中分析对由文献挑选出的因素合理性，并根据问卷数据结果构建难度模型。第四步，从最终得出因素水平的新定义和难度模型中，分析优缺点。

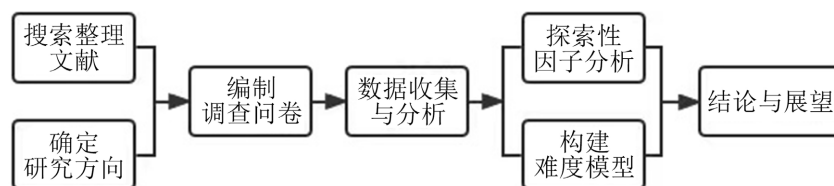


Figure 1. Research framework

图 1. 研究框架

## 2.2. 问卷编制

在鲍建生等教授五因素模型的基础上，结合专家意见，将试题考查的影响因素确定为知识点数量、情境、运算、推理、认知、开放性、隐蔽性、梯度、未知参数数量等 9 个影响因素，各因素水平被平均分为 5 个层次，由低到高分别赋值为 1、2、3、4、5。《初中数学试题难度影响因素调查问卷》的编制选取了宁波市近八年中考试卷中具有代表性且已知难度系数的中考题，在明确告知被调查者 9 个试题难度因素的判断标准后，请被调查者根据所给题目以及参考答案分别从 9 个因素进行评判打分。问卷共 6 题，问卷调查细目表见表 1。

Table 1. Questionnaire breakdown

表 1. 问卷调查细目表

题号	已知难度系数	考查的影响因素
1	0.93	
2	0.92	
3	0.71	知识点数量、情境、运算、推理、认知、 开放性、隐蔽性、梯度、未知参数数量
4	0.71	
5	0.73	
6	0.59	

下面以问卷第 4 题为例，具体说明试题考查的 9 个影响因素的水平判断及赋分方式。问卷第 4 题(见图 2)选自 2016 年宁波市中考第 23 题[7]，题目本身稍有难度，也不存在赋值为 0 的因素，相较于其他例题，此题更适合当作例题进行具体分析、向读者展示赋分方式。

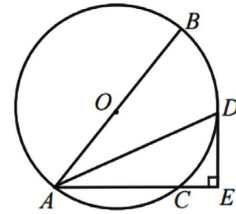
表 2 展示了具体的 9 个影响因素的赋分表。对照表格各行对本题依次进行分析。本题涉及直线与圆的位置关系、矩形的判定与性质、勾股定理、角平分线的性质、等腰三角形的性质等知识点，故知识点数量为 5 个及 5 个以上知识点，赋值 5；本题题干只运用了数学语言和相关文字叙述，故情境为数学情境，赋值 1；本题第一小题无运算，第二小题分别计算 AF、OF 和 DE，故运算为三步以上数值运算，赋值 3；本题第一小题由角平分线的性质和等腰三角形的性质推出直线平行，为 1 步推理，第二小题由勾股定理计算线段长度，再根据矩形的性质求得线段长度，有 2 步推理，故本题共为 3 步推理，赋值 4；本题是计算与推理结合的综合性问题，需要学生逐步推理，进行转化，故是分析水平，赋值 4；本题解

题方法唯一, 故开放性赋值 1; 本题隐含条件为角平分线的性质、等腰三角形的性质、勾股定理和矩形的判定和性质等, 故有四个隐含条件, 赋值 5; 本题仅第二小题要用到第一小题的结论, 故梯度赋值 3; 本题为纯数值运算且不需要设未知数, 故未知参数数量赋值 1。

4.如图, 已知 $\odot O$ 的直径 $AB=10$ , 弦 $AC=6$ ,  $\angle BAC$ 的平分线交 $\odot O$ 于点 $D$ , 过点 $D$ 作 $DE \perp AC$ 交 $AC$ 的延长线于点 $E$ .

(1) 求证:  $DE$  是 $\odot O$ 的切线.

(2) 求 $DE$ 的长.



**参考答案:** (1) 连结 $OD$ ,

$\because AD$  平分 $\angle BAC$ ,  
 $\therefore \angle DAE = \angle DAB$ ,  
 $\because OA = OD$ ,  $\therefore \angle ODA = \angle DAO$ ,  
 $\therefore \angle ODA = \angle DAE$ ,  
 $\therefore OD \parallel AE$ .  
 $\because DE \perp AC$ ,  
 $\therefore OD \perp DE$ .  
 $\therefore DE$  是 $\odot O$ 的切线.

(2) 过点 $O$ 作 $OF \perp AC$ 于点 $F$ ,

$\therefore AF = CF = 3$ ,  
 $\therefore OF = \sqrt{AO^2 - AF^2} = \sqrt{5^2 - 3^2} = 4$ .  
 $\because \angle OFE = \angle DEF = \angle ODE = 90^\circ$ ,  
 $\therefore$  四边形 $OFED$ 是矩形.  
 $\therefore DE = OF = 4$ .

Figure 2. Question 4 of the questionnaire

图 2. 问卷第 4 题

Table 2. Impact factor scoring scale

表 2. 影响因素赋分表

影响因素	水平				
	1	2	3	4	5
知识点数量	1 个知识点	2 个知识点	3 个知识点	4 个知识点	5 个及 5 个以上知识点
情境	数学情境	个人生活	公共常识	科学常识	专业情境, 如金融
运算	无运算	3 步以下数值运算	3 步以上数值运算	3 步以下符号运算	3 步以上复杂符号运算
推理	无推理	1 步推理	2 步推理	3 步推理	4 步及 4 步以上推理
认知	识记(背诵知识、公式)	理解(掌握知识意义)	运用(用已理解知识解决新问题)	分析(逐步推理、转化问题)	评估与创造(未接触过或无现成解决方案, 创造性地解决问题)
开放性	有唯一的正确答案和解法	有唯一解法且正确答案有 2 种	有唯一解法且正确答案有 3 种	有唯一解法且正确答案有 4 种	有多种解法且正确答案唯一
隐蔽性	条件直接	有 1 个隐含条件	有 2 个隐含条件	有 3 个隐含条件	有 4 个隐含条件
梯度	无小题	有小题但各小题间相互独立	有 1 个小题用到上一小题的结论	有 2 个小题用到前面小题的结论	有 3 或 3 个以上小题用到前面小题的结论
未知参数数量	不含未知参数	含有 1 个未知参数	含有 2 个未知参数	含有 3 个未知参数	含有 4 个及 4 个以上未知参数

为了提高调查对象回答问题的准确性和客观性, 本次调查的主要对象是责任心较强、教学研究水平较高、且对初中数学试题及其难度有较为客观和深入认识的数学师范专业在读本科生。在发放问卷之前,

调查对象已对各因素各水平判断标准有非常明确的认识。调查共发放 126 份问卷, 回收问卷 126 份, 回收率 100%, 回收后剔除了基本信息不全、未作答、数据缺失严重等的问卷 5 份, 故有效问卷共 121 份, 有效率 96%。

信度是表明测试可靠性的指标, 其实是检测测试卷的一致性、稳定性。一般情况下, 若调查问卷的信度系数达到 0.9 以上, 说明该调查问卷的信度较高; 信度系数在 0.8 以上, 说明信度较好, 如果信度系数低于 0.5, 则此调查问卷调查结果的可行性比较差, 不建议采纳该调查问卷数据。为了检验此次问卷的信度, 利用 SPSS 进行统计分析, 结果如表 3。

**Table 3.** Questionnaire reliability checklist  
**表 3.** 问卷信度检验表

克隆巴赫 Alpha	项数
0.710	54

经检验, 问卷的克隆巴赫信度系数  $\alpha = 0.710$ , 结果表明问卷具有可接受的内部一致性。

### 3. 结果分析

下面对问卷数据进行探索性因子分析, 将难度界定为由各因素线性组合表示的量化值。

#### 3.1. 初中数学试题难度影响因素的探索性因子分析

数据运行结果显示, 初中数学试题难度影响因素其中 8 个因素在 0.05 显著性水平都相关。只有开放性因素与知识点数量、推理、认知与隐蔽性因素的相关性(分别为 0.116、0.250、0.305 和 0.149)未达 0.05 显著性水平下的中等相关即 0.4 的要求, 故认为两者没有显著线性关系的特征。结合旋转成分矩阵中开放性在抽取出的三个因子中的占比(分别为 0.077、0.437 和 0.476)均未超过 0.5, 因此考虑先将开放性因素删除。

再进行 KMO 与 Bartlett 检验 9 个影响因素, 其结果 KMO 值为 0.734, 说明各影响因素间存在共同因素, 适用于信度分析, 即用主成分分析法提取问卷中的公共要素。

综合考虑上述各影响因素的相关性及抽取因子上的负荷量差异等指标后, 将开放性因素删除, 初中数学试题难度影响因素的探索性因子分析至此结束, 一共删除了 1 个影响因素即开放性因素, 剩余 8 个影响因素。

**Table 4.** Rotated component matrix  
**表 4.** 旋转成分矩

	成份		
	1	2	3
推理	0.810	0.152	0.110
知识点数量	0.781	-0.162	0.111
隐蔽性	0.705	-0.001	0.032
认知	0.610	0.344	0.226
未知参数数量	-0.023	0.863	-0.001
运算	0.537	0.654	-0.078
梯度	0.201	-0.092	0.856
情境	-0.366	0.463	0.530

对 8 个影响因素进行效度分析,从表 4 可知各影响因素都在某一成份上的负荷量相对较高,从表 5 可知三个成份累积的解释的总方差为 65.731%,即这三个影响因素能够较好的表示这 8 个指标。重新进行 KMO 与 Bartlett 检验,其 KMO 值提高为 0.826,说明影响因素间存在共同因素,且再次说明删除开放性的必要性。从整体上看,各个测量项均呈现出 0.001 水平的显著性,即因子与测量项之间有着良好的对应关系,聚合效度较好。重新进行信度检验,得克隆巴赫信度系数  $\alpha$  为 0.844,说明内部一致性均较好。从效度分析结果可知,这五个因素之间存在相关关系,信效度较高,适合建立模型,且应以成份 1、成份 2、成份 3 这三个因素为主要分析因素。根据结果,将成份 1、2、3 分别命名为概念性要素、程序性要素、情境性及层次性要素。其中概念性要素包括推理、知识点数量、隐蔽性、认知;程序性要素包括未知参数数量和运算;情境性及层次性要素包括梯度和情境。

**Table 5.** Total variance explained

**表 5** 解释的总方差

成份	初始特征值			提取平方和载入			旋转平方和载入	
	合计	方差的%	累积%	合计	方差的%	累积%	合计	方差的%
1	2.717	33.968	33.968	2.717	33.968	33.968	2.599	32.489
2	1.494	18.671	52.639	1.494	18.671	52.639	1.564	19.546
3	1.047	13.093	65.731	1.047	13.093	65.731	1.096	13.697
4	0.761	9.515	75.246					
5	0.642	8.023	83.270					
6	0.529	6.614	89.883					
7	0.419	5.242	95.125					
8	0.390	4.875	100.000					

### 3.2. 初中数学试题难度模型建立

#### 3.2.1. 模型建立

**Table 6.** Regression coefficient table

**表 6.** 回归系数表

模型	非标准化系数		标准系数	t	Sig.
	B	标准 误差	试用版		
(常量)	-0.103	0.010		-10.682	0.000
知识点数量	0.027	0.002	0.267	13.549	0.000
情境	0.001	0.002	0.010	0.612	0.541
运算	0.040	0.003	0.350	15.826	0.000
推理	0.011	0.002	0.123	5.890	0.000
认知	0.029	0.003	0.200	10.214	0.000
隐蔽性	0.008	0.002	0.071	3.771	0.000
梯度	-0.016	0.002	-0.105	-6.814	0.000
未知参数数量	0.025	0.002	0.244	12.413	0.000

由前面探索性因子分析得出三个成分要素，可观察出影响因素之间的相关性，接着对影响因素与难度建立一元线性回归方程，期望观察出各个影响因素对难度的影响程度，并对难度预测提供借鉴案例。通过回归分析得到表 6。

由表 6 可知，将上述 8 个影响因素与难度之间的量化关系用数学线性表达式表示为

$$N = 0.267A + 0.01B + 0.35C + 0.123D + 0.2E + 0.071F - 0.105G + 0.244H$$

其中， $N$  表示初中试题难度， $A$  表示知识点数量、 $B$  表示情境、 $C$  表示运算， $D$  表示推理、 $E$  表示认知， $F$  表示隐蔽性、 $G$  表示梯度， $H$  表示未知参数数量。具体运用于初中数学试题难度量化分析时，因素已经标准化且每个影响因素的量化值范围为 1~5 之间，故最终线性表达式得出的难度量化范围也在 1~5 之间。

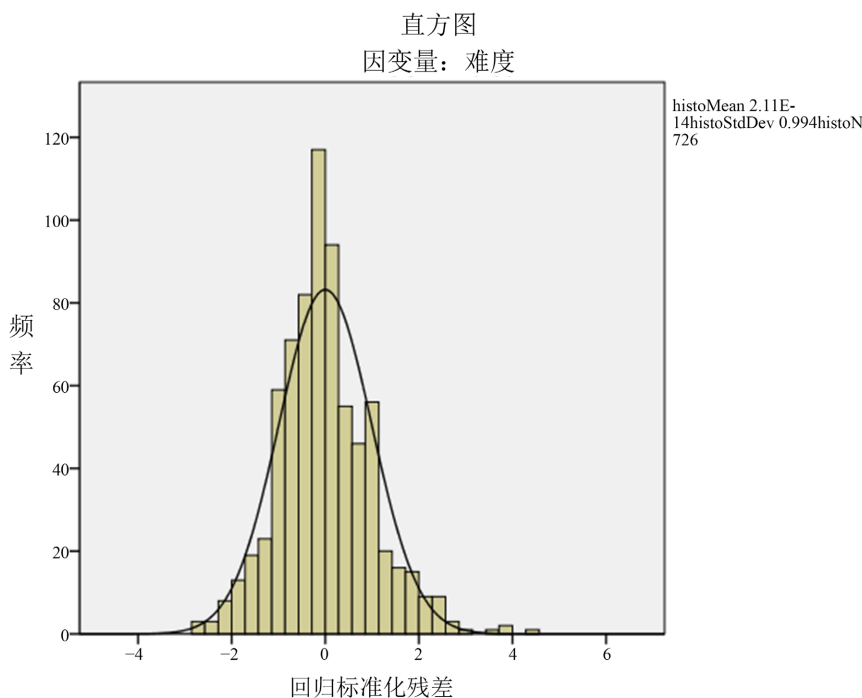
另外，上述数据处理过程中会造成各影响因素对初中数学试题难度的难度系数与实际统计分析所得难度系数不一致，但两者所反映的初中数学试题难度整体趋势是一致的[8]。

### 3.2.2. 模型检验

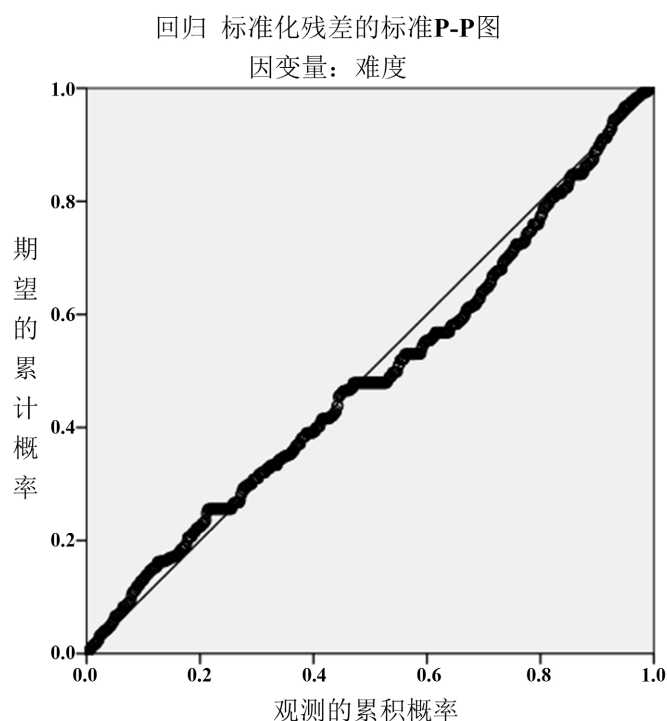
**Table 7.** Model fit test  
**表 7.** 模型拟合度检验

模型	R	R 方	调整 R 方	标准估计的误差
	0.921	0.849	0.847	0.04769

由表 7 可知，第二列 R 是复相关系数，其表示自变量和因变量之间的密切程度。第三列 R 方是复相关系数的平方，又称决定系数。两者值越大越好，其中观察到 R 方为 0.849，大于 0.8，可知模型拟合度好[9]。



**Figure 3.** Histogram of standardized residuals  
**图 3.** 标准化残差直方图



**Figure 4.** Standard P-P diagram  
**图 4.** 标准 P-P 图

观察图 3，可知标准化残差具有正态分布的趋势；由图 4 可知，散点基本接近直线，说明残差正态性结果好。因此，模型数据拟合度很好。

### 3.2.3. 模型分析

通过回归模型各个影响因素的系数可以知道其对难度的影响大小、正负影响等。由于运算的回归系数为 0.350，是绝对值最大的回归系数，故运算对难度的影响最大，而绝对值最小的情境对难度的影响相应是最小的。同时，只有梯度的回归系数是负数，故梯度对难度是负影响，其他因素对难度都是正影响，即梯度赋值越大，难度反而越小。

## 4. 研究结果

基于以上研究，可得本文研究结果如下：

经探索性因子分析、KMO 与 Bartlett 检验，在知识点数量、情境、运算、推理、认知、开放性、隐蔽性、梯度和未知参数数量 9 个因素中，可将开放性这一因素删除，最终确定 8 个因素。

对 8 个影响因素进行信度分析，将推理、知识点数量、隐蔽性和认知命名为概念性因素，未知参数数量和运算命名为程序性要素，梯度和情境命名为情境及层次性要素。

以试题难度为因变量，知识点数量、情境、运算、推理、认知、隐蔽性、梯度和未知参数数量为自变量，建立一元线性回归方程：

$$N = 0.267A + 0.01B + 0.35C + 0.123D + 0.2E + 0.071F - 0.105G + 0.244H$$

其中  $N$  表示初中试题难度， $A$  表示知识点数量， $B$  表示情境， $C$  表示运算， $D$  表示推理， $E$  表示认知，表示隐蔽性， $F$  表示梯度， $G$  表示未知参数数量。

分析一阶线性方程，可知运算与知识点数量对难度影响较大，情境和隐蔽性对难度的影响较小，除



梯度外, 其他因素与难度均呈正相关。

## 5. 总结

由探索性因子分析可知, 数学试题难度绝大部分影响因素彼此具有较高的相关性。因此, 教师在命题时需综合考虑多种难度影响因素。同时, 本文建立的数学试题难度模型在一定程度上反映了各因素对于试题难度的影响差异。对于造成试题难度的主、次因素的分析, 有利于教师在试题编写时, 有针对性地设置试题考察点。例如, 由综合难度系数可得运算与知识点数量相较于其他因素对试题难度影响较大。换句话说, 增加运算与知识点数量会大幅增加试题难度。“情境”与“隐蔽性”因素承载的难度较小。该情况的出现可能由于“情境”维度关于“公共常识、科学常识、专业情境”水平方面存在明显的不足。这也折射出对新课改所提倡的“情境”等内容的重视程度偏低, 但丰富试题的情境性有利于学生增强应用意识。由此, 针对命题及教学, 教师应当巧设试题背景, 重视社会现实问题, 与时俱进, 使学生深刻体会到数学与生活的密切联系, 从而更好地发挥情境在引导素质教育、促进学生全面发展的作用。

本文构造的数学试题绝对难度模型突破了以往试题估计的主观性和局限性, 在事先预测试题难度和试题的量化分析处理方面具有较大的优势, 但该模型也存在不足之处。首先, 本文虽利用探索性因子分析对各因素内部间的相关性做了说明, 删除了相关性较小因素。但通过文献借鉴与主观筛选确立的难度影响因素可能并非全面; 其次, 各影响因素的不同水平之间的权重仅简单使用 1、2、3、4、5 数字进行赋值, 它们之间的等距性存在问题, 后期可以通过 AHP 层次分析理论进一步地改进, 从而使结果更加合理。因此, 要完善该数学试题难度模型, 后续仍有许多工作。

## 基金项目

浙江省高等教育学会 2023 年度高等教育研究课题(KT2023190)。

## 参考文献

- [1] 邵志芳, 刘永芳, 钟毅平. 关于问题难度的实验研究[J]. 心理科学, 1996(5): 278-281+320.
- [2] 徐建乐. 数学试题难度的指标体系及量化[J]. 中学数学教学, 1996(3): 9-12.
- [3] 杜明荣. 高中物理试题难度的影响因素研究[D]: [硕士学位论文]. 重庆: 西南大学, 2008.
- [4] 王建磐, 鲍建生. 高中数学教材中例题的综合难度的国际比较[J]. 全球教育展望, 2014, 43(8): 101-110.
- [5] 武小鹏, 张怡. 中国和韩国高考数学试题综合难度比较研究[J]. 数学教育学报, 2018, 27(3): 19-24+29.
- [6] 李保臻, 石焯. 中国大陆与台湾地区高考数学试题难度比较研究——以 2016-2018 年大陆全国卷 I 与台湾指考试题为例[J]. 数学教育学报, 2020, 29(1): 58-64.
- [7] 宁波市教育考试院. 宁波市 2016 年初中学业水平考试试题解析[EB/OL]. <http://nbcea.nbedu.net.cn/school/NeiRead.aspx?id=552>, 2017-03-22.
- [8] 蔡庆有, 邝孔秀, 宋乃庆. 小学数学教材难度模型研究[J]. 教育学报, 2013, 9(5): 97-105.
- [9] 孙佳. 初中生数学问题提出水平及影响因素研究[D]: [硕士学位论文]. 北京: 中央民族大学, 2021.