

基于 Rasch 模型的初中数学模型观念的测评研究

马文静

扬州大学数学科学学院 江苏扬州

【摘要】随着核心素养教育理念的推进，数学建模能力成为初中数学教育的重要目标。本研究以 Rasch 模型为理论框架，构建初中数学模型观念评价指标体系，编制测试工具，并对某市某中学初二学生进行实证分析。研究发现：学生模型观念整体水平较好，但运用能力较弱；中等水平学生占比较高；测试卷信度与区分度符合测量要求。基于此，提出强化建模教学目标、分层教学、增加实践应用等教学建议，为数学核心素养培养提供参考。

【关键词】Rasch 模型；初中数学；模型观念；核心素养

【收稿日期】2025 年 2 月 18 日 **【出刊日期】**2025 年 3 月 18 日 **【DOI】**10.12208/j.aam.20250007

Elementary school mathematical model concept based on the Rasch model Assessment research

Wenjing Ma

School of Mathematical Sciences, Yangzhou University, Yangzhou, Jiangsu

【Abstract】With the advancement of the concept of core literacy education, mathematical modeling ability has become an important goal of junior high school mathematics education. Based on the Rasch model as a theoretical framework, this study constructed an evaluation index system of junior high school mathematical model concepts, compiled test tools, and conducted an empirical analysis on the second year junior high school students in Changzhou City. The results show that the overall level of students' model concept is good, but the application ability is weak. The proportion of students with intermediate level is relatively high; The reliability and discrimination of the test volume met the measurement requirements. Based on this, some teaching suggestions are put forward, such as strengthening the teaching objectives of modeling, hierarchical teaching, and increasing practical application, so as to provide reference for the cultivation of mathematics core literacy.

【Keywords】Rasch model; Middle school mathematics; Model concept; Core literacy

1 问题提出

《义务教育数学课程标准（2022 版）》明确提出将模型观念作为数学核心素养之一^[1]。初中阶段是学生抽象思维发展的关键期，需通过数学建模连接理论与实际问题^[6]。然而，传统评价方法依赖经典测量理论，存在样本依赖性强、无法精准评估个体能力等问题。项目反应理论（IRT）及 Rasch 模型的应用，为教育评价提供了更科学的工具。本研究旨在通过 Rasch 模型评估学生模型观念水平，探索培养策略^[7]。

2 理论基础

2.1 Rasch 模型

Rasch 模型基于 IRT 理论，通过概率模型评估个体潜在能力，具有单维性、题目难度与能力参数独立于样本等优势^[2]。本研究采用 Winsteps 软件进行数据分析，通过拟合度、信度、区分度等指标验证测试工具的有效性。

2.2 数学核心素养框架

结合布鲁姆教育目标分类学中的认知层次目标^[3]（记忆、理解、应用、分析、综合、评价）与 SOLO 分类理论对认知水平层次的划分^[4]（前结构层次、单点结构层次、多点结构层次、关联结构层次和抽象扩展结

构层次），构建数学模型观念三级能力：认识数学建模能力、建立数学建模能力、运用数学建模能力。

3 研究过程与结果

3.1 测试工具与样本

基于 Rasch 模型，编制覆盖“认识、建立、运用”三维度的数学模型观念测试卷（共 13 题），题目难度范围-1.3~1.45 Logit。选取常州市某中学初二 113 名学生为样本（男生 54%，女生 46%），通过 45 分钟闭卷测试收集数据，采用 Winsteps 软件进行模型拟合分析。

3.2 研究结果

在本项研究中，我们采用了 Winsteps 软件来分析参与者的得分数据，并据此构建了本次测试卷的 Rasch 模型^[5]。该模型的构成要素包括：全面的质量检验图表、标准残差图表（用于单维性检验）、各个题目的拟合统计表（分析题目的各维度拟合情况）、以及怀特图（用于分析参与者的能力和题目的难度），这些工具共同用于综合评估测试卷的设计质量。接下来，我们将对模型中的每个组成部分进行详尽的讨论。

3.2.1 测试卷整体质量检验表

表 1 测试卷整体质量检验表

Measure	Infit		Outfits		Separation	Reliability
	MNSQ	STDZ	MNSQ	STDZ		
Person	0.96	1.00	0.10	1.00	0.10	0.00
Item	0.00	1.00	0.10	1.00	0.00	0.87

表 1 展示了 MNSQ 和 ZSTD 指标，这些指标是评估测试题目与 Rasch 模型 拟合度的关键工具。MNSQ 指标的合理区间为 0.5 至 1.5，而 ZSTD 指标的接受区间则为-2 至 2。理想状态下，MNSQ 值应尽可能地接近 1，ZSTD 值则应趋近于 0，这表明题目与模型之间的拟合度较高。根据表 4-5 的数据，无论是参与者的内在还是外在 MNSQ 值，还是题目的相应 MNSQ 值，均达到 1.00。同时，参与者和题目的 Infit ZSTD 值均为 0，参与者的 Outfit ZSTD 值为 0.1，题目的为 0，这些数据均反映出测试题目与模型之间达到了高度拟合。在评估测试卷的质量时，信度是一个重要的指标，它反映了测试结果的稳定性和可靠性。信度值介于 0 到 1 之间，通常情况下，高于 0.8 的信度值意味着测试具有较高的可信性，本测试的信度值为 0.87，显示出测试卷在评估上的稳定性和可靠性。区分度是衡量测试卷对不同能力水平参与者区分能力的另一个重要指标，其值通常要求大于 2。测试卷的区分度为 2.61，这表明测试卷能够有效区分不同水平的参与者。最后，Measure 值提供了 Rasch 模型对测试题目难度和参与者能力的估计平均分，本研究中参与者能力的估计平均分值为 0.96，这为理解测试卷整体难度和参与者能力水平提供了重要参考。

3.2.2 测试卷项目标准残差图

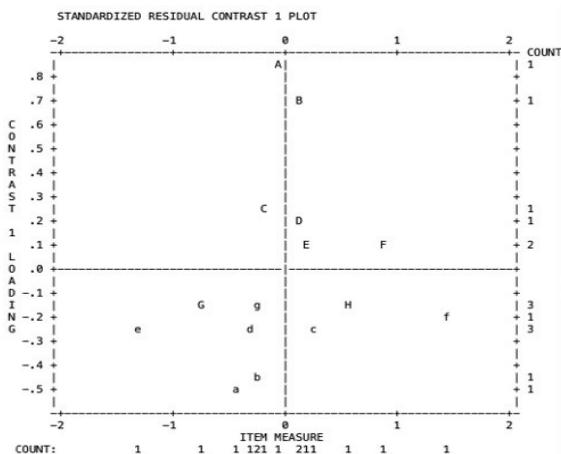


图 1 测试卷项目标准残差插图

若某一题目的单维性检验指标值位于 [-0.4, +0.4] 的范围内，则该题目被认为符合单维性要求，意味着参与者的回答主要受到单一因素的影响。测试题目的单维性检验结果通常以标准残差图的形式展示，其中不同的大小写字母代表不同的题目项目。在图 1 中，展示了 15 道题目的单维性检验结果。根据标准残差图，有 11 道题目达到了单维性标准，这些题目分别用 C、D、E、F、G、H、c、d、e、f、g 表示。然而，题目 A、B、a、b 的指标值超出了这一区间，对应于题目 1 (a)、1 (b)、3、8，这暗示学生对这些题目的回答可能受到数学模型观念之外的其他因素的干扰。总体而言，由于绝大多数题目显示出单维性特征，这证实了测试题目设计的有效性，并能够精确地衡量学生的数学模型观念能力。

3.2.3 测试卷项目各维度拟合图

表 2 测试卷项目各维度拟合图

Measure	Infit		Outfits		PT-Measure CORR	Item
	MNSQ	STDZ	MNSQ	STDZ		
1.45	1.01	0.20	1.02	0.20	0.26	9 (b)
0.91	1.14	2.90	1.15	2.50	0.05	10
0.56	1.05	1.00	1.03	0.50	0.18	11 (b)
0.23	0.93	-0.90	0.90	-1.10	0.35	9 (a)
0.19	0.99	-0.10	0.99	-0.10	0.26	4 (b)
0.10	0.88	-1.40	0.84	-1.60	0.42	1 (a)
0.10	1.01	0.20	0.97	-0.20	0.23	11 (a)
-0.04	0.96	-0.40	0.94	-0.40	0.29	1 (b)
-0.18	0.99	-0.10	1.00	0.00	0.23	6
-0.23	0.95	-0.40	0.89	-0.70	0.31	3
-0.28	0.96	-0.30	0.97	-0.20	0.27	7
-0.33	1.03	0.30	1.03	0.20	0.16	2
-0.44	1.06	0.50	1.05	0.40	0.12	8
-0.74	1.01	0.10	1.08	0.50	0.15	4 (a)
-1.30	1.02	0.20	1.07	0.40	0.12	5

本研究利用 Winsteps 软件生成了测试卷各题目拟合统计表，具体内容见表 2。根据表 2 的数据可知，各题目难度值介于 [-1.3, 1.45] 之间，分布范围广，能够在较大范围内对被试能力水平进行评估。试题的 Outfit MNSQ 数值均在 [0.5, 1.5]，表明试题与被试水平拟合成功。点-测量相关系数（PT-Measure CORR）可以检验学生在题目的得分与测试卷总得分的一致性程度，其理想范围是 0-1 之间，由表 2 可知，各题目点-测量相关系数在合理范围内，表明各题目与测试卷的相关性较高。

3.2.4 被试能力与试题难度匹配度分析

怀特图是一种评估项目难度与考生能力匹配度的工具。该图的左侧采用 Logit 量表，从低到高依次排列考生的能力与题目难度。图 2 中的虚线表示 Logit 尺度，其中 M 代表中等水平，S 代表 1 个 Logit 单位，T 代表 2 个 Logit 单位。量表左侧显示考生能力分布，每个“#”符号代表一名考生，而右侧则展示各题目的难度分布。随着量表向上，考生的能力逐渐增强，同时题目难度也相应提高。本研究利用 Winsteps 软件绘制了试卷项目与考生的怀特图。通过分析图 2，可以发现考生能力的平均值高于题目的难度平均值，这表明试卷对考生来说相对容易；考生的能力分布介于 -1Logit 至 3Logit 之间，并且大致呈正态分布，表明中等能力的考生数量较多；题目难度的分布范围在 -2Logit 至 2Logit 之间，大多数题目的难度分布较为

合理；然而也存在一些高能力考生没有合适的题目与之匹配，这表明试卷的难度范围未能覆盖所有考生的能力水平。

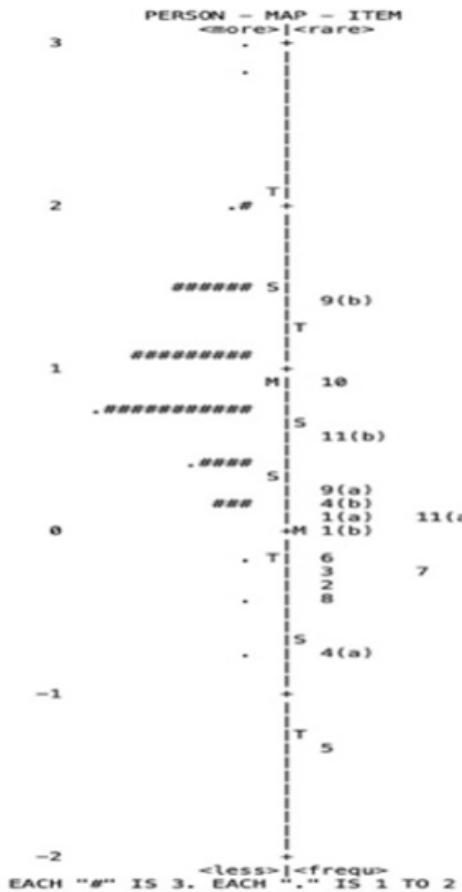


图 2 被试测试卷项目-被试怀特图

4 研究结论

4.1 被测试学生的总体模型观念水平较好

通过研究可得出被测试学生的模型观念水平较好，但这也与测试卷难度较小有关。从表 1 中用 Measure 值测出参与者能力估计平均分为 0.96，整体上我们可以看出被测试的学生总体质量是较好的，有着较好的模型观念水平。

4.2 学生的模型观念水平受多因素影响

从项目标准残差图可看出，部分测试题不满足单维性，说明这些测试题不仅受到学生模型观念水平的影响，还可能会受到其他因素的影响。但是绝大多数的题目仍是呈现出单维性特征的，这说明了测试题目的有效性且能精确地衡量学生的数学模型观念能力。

4.3 学生模型观念水平处于中等水平的人数较多

测试卷整体题目难度分布范围较广，处于中等水平的学生较多。从测试卷项目各维度拟合情况表来看，各题目难度值介于 [-1.3, 1.45] 之间，结合怀特图，我们发现测试卷能够向较大范围使用。

4.4 被测试学生在运用数学模型能力方面有所欠缺

由怀特图可看出，考察学生运用数学模型能力的题目有五题，其中第十题、第十一题的第一小问和第二小问这三道题，学生的作答情况不是很好，而相较于认识数学模型能力和建立数学模型能力仅有一到两题学生答题情况不佳。

5 教学建议

5.1 教师要强化学生建立模型观念的教学目标

模型观念作为初中数学要求的核心素养之一，教师在数学教学中应将培养模型观念作为教学的重要目标，在授课、讲题等教学活动中贯穿培养^[8]。同时，鉴于学生在数学模型观念方面表现良好，教师应继续在课程中强化模型观念的教学，帮助学生进一步发展和深化这一能力。

5.2 教师可以增加实践应用提高学生模型观念

教师要鼓励学生将数学模型应用于现实生活中的问题，通过实际案例分析和项目式学习，增强学生对数学模型实用性的认识和应用能力。

5.3 教师可采用分层教学活动，注意学生个体差异

由于学生能力分布广泛，建议教师设计不同难度层次的培养数学模型观念的教学活动，以满足不同水平学生的需求，特别是为高水平学生提供更具挑战性的问题解决机会。而对于低层次水平的学生，教师也应该注重学生的基础知识和模型观念设计个性化的学习路径，为每个学生提供适合其能力和兴趣的学习材料和挑战。

5.4 教师可以促进学生高阶思维技能

研究结论得出模型观念处于中等水平的学生较多，针对这一问题，教师应该在教学中融入更多需要高阶思维技能的活动如创新、批判性思维和问题解决，以培养学生在数学模型建立和应用方面的高级能力。另外，教师也要利用计算机软件、在线资源和其他教育技术工具，增强教学的互动性和学生的参与度，同时为学生提供模拟和实验数学模型的机会。

参考文献

- [1] 中华人民共和国教育部. 义务教育数学课程标准 (2022 年版)[S]. 北京: 北京师范大学出版社, 2022.
- [2] 陈萍萍, 程晓堂. 基于 Rasch 模型的课程标准参照性测试任务效度研究 [J]. 外语教学, 2023,44(06): 56-62.
- [3] 丁念金. 霍恩斯坦教育目标分类与布卢姆教育目标分类的比较 [J]. 外国教育研究, 2004,(12): 10-13.
- [4] 闫明智, 鄢银银. 基于 SOLO 分类理论的高中生区域认知思维水平评价 [J]. 地理教育, 2021,(05): 4-7.
- [5] 叶杨. 基于 Rasch 模型的初中生数学推理能力评价研究 [D]. 江南大学, 2023.
- [6] 丁芸. 在初中数学教学中培养学生的模型思维的探索[J]. 新智慧, 2019,(32): 127-128.
- [7] 鲍建生, 章建跃. 数学核心素养在初中阶段的主要表现之七: 模型观念[J]. 中国数学教育, 2022,(23):3-8.
- [8] 喻平. 发展学生数学核心素养的一个教学模式建构 [J]. 数学通报, 2023, 62 (09): 1-6+11.

版权声明：©2025 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

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



OPEN ACCESS