

证据驱动的青少年数学学习力测量模型构建研究

隋元

上海诺美高级中学 上海 201100

【摘要】：随着教育评价范式从单一成绩评定向综合能力评估转变，青少年数学学习力的科学测量成为教育心理学的重要课题。本研究旨在构建一个证据驱动的、多维度融合的青少年数学学习力测量模型，并基于实测数据对模型进行初步验证与分析。研究通过跨学科方法，整合数学应试能力、数学问题解决素养（数学素养）及非认知因素（数学焦虑、自我效能感、兴趣等），构建了一个综合评估模型。本研究对象为通过上海某教育自媒体招募的上海市1-9年级中小学生及其主要养育者。研究分为两个阶段：模型构建阶段通过文献分析与前期调研（ $n=75$ 组家庭）确立模型维度与测量工具；模型实测阶段则随机选取其中20组家庭进行一对一的综合测评。测评工具包括自编的标准化数学能力试卷、情境化问题解决实验，以及改编的数学焦虑、自我效能感与创造力量表，并辅以对养育者的结构化访谈。研究发现，在主动寻求测评的家庭中，母亲（占比89.3%）及女生（占比76%）占绝大多数，提示特定群体对数学学习存在更高的关注或焦虑。主要养育者对孩子数学能力的评估普遍低于其实际表现，两者呈显著正相关，即孩子实际成绩越好，家长的低估倾向越明显。数学兴趣、数学素养与实际成绩三者间均存在正向关联，其中数学兴趣与数学素养之间的相关性最强。本研究构建的模型不仅为理解青少年数学学习力提供了多维视角，其测评结果也可为学生个性化发展、家长教育决策及教师教学干预提供实证依据。未来研究将在扩大样本量的基础上，进一步优化模型，并探索其在不同学段与文化背景下的适用性。

【关键词】：数学学习力；测量模型；证据驱动；非认知因素；数学素养；教育评估

DOI:10.12417/2705-1358.26.14.043

1 研究问题的兴起

在教育心理学和学业成就研究领域，青少年的数学学习力一直是备受关注的主题。随着教育评价的重心逐渐从简单的成绩评定转向学生综合能力的评估，加之非认知因素在数学学习中的重要影响在当今时代背景下尤为显著，如何科学地构建并准确测量学生的数学学习力成为了一个亟待解决的问题，在笔者的一线经验中也感触颇深。在这样的背景下，“青少年数学学习力测量模型构建”作为一个新兴的研究主题，意在通过跨学科的方法，结合定量与定性分析，深入探究影响青少年数学学习力的关键因素，并构建一个既包含数学应试能力测量，又融入实际数学认知能力素养、情感态度的综合模型。这一模型的建立不仅有助于帮助学生更全面地理解数学学习过程，而且能为家长及学生本人提供学生的实际数学能力参考，为教师提供更为精确的教学指导，为政策制定者提供科学的决策支持，从而促进每个学生的个性化发展。此研究问题的兴起反映了当前教育界对提升教学质量、实现真正意义上的教育过程公平及培养未来创新人才的迫切需求。通过证据驱动的研究方法，可以确保研究成果的实证基础，增强结论的有效性，并为教育实践提供切实可行的改进建议。因此，该研究议题不仅具有重要的理论价值，也具备显著的实践意义。

2 研究问题

本研究聚焦于2个部分，即青少年数学学习力模型构建以及基于该模型以青少年及主要养育者作为被试的综合评估数据分析。青少年数学学习力模型构建主要通过前期文献查阅积累理论基础，在理论基础之上先通过简单调研结合笔者在中小学工作的经验，来确定模型的相应维度及测量方式，模型初步构建后，通过量表、实验、观察、访谈等形式对每一个被试个体及其主要养育者进行实测，对每个个体情况作记录并给出测评报告，并基于整体数据进行变量间的相关分析。

3 文献综述

在教育研究领域，青少年的数学学习力是衡量学生数学素养和应对数学挑战能力的关键指标。近年来，随着认知科学、心理测量学和数据科学技术的发展，研究者开始尝试构建更为精确和全面的数学学习力模型，并通过实证研究进行验证和细化。一系列的文献对数学学习力的多维度特性进行了深入探讨，如问题解决能力、逻辑推理、数学态度以及自我效能感等^[1-2]。^[1]强调了问题解决过程中认知策略的重要性，并提出了一个包含知识、策略和信念系统的数学问题解决框架。^[2]则关注了教学实践对学生数学态度的影响，认为积极的数学态度有

助于提升学习力。近年来,国内不少研究也开始聚焦于非认知因素对数学学业成就的影响。^[7]认为,数学焦虑水平能显著负向预测数学成绩,而数学自我效能感能显著正向预测数学成绩。而随着AI时代的来临,对于数学能力的评估也将视线转移到了数学创造力上。^[8]发现,数学焦虑对数学创造力有负向预测作用,数学自我效能感在数学焦虑与数学创造力之间起全中介作用,即数学焦虑通过数学自我效能感影响数学创造力。另一方面,实证研究方法的进步为模型的检验提供了新工具。例如,利用项目反应理论(IRT)和结构方程模型(SEM)等高级统计技术可以更准确地量化和分析学生的数学能力^[3]。证据驱动的教育研究逐渐成为推动该领域发展的关键力量。通过大规模数据集的分析,研究者能够更全面地理解影响数学学习力的各种因素,并为教育干预提供指导。例如,^[4]通过分析国家教育进展评估(NAEP)的数据,探讨了课程、教学和学生背景如何共同作用于学生的数学成就。然而,现有文献也指出了研究中存在的问题,比如模型的文化适应性、测量工具的普适性以及数据驱动研究可能存在的选择偏差等问题(e.g.,Baker et al.,2014)。此外,随着教育技术的发展,数字游戏、在线学习平台等新型学习工具对于数学学习力的影响也开始受到研究者的关注^[5]。综上所述,青少年数学学习力测评模型的构建与实测是一个跨学科、多方法的综合研究领域,涉及心理学、教育学、统计学等多个领域的知识。现有研究普遍聚焦于教学与数学成就之间的关系,对于数学学科所需底层素养能力的关注与测量较少,且证据驱动研究较少,现有研究已关注到一些非认知因素维度的影响,但尚未就此给出结合我国学情青少年数学学习力测评模型。

4 研究方法

本节研究方法将分为以下几个部分展开,即研究对象、研究工具、数据收集与处理。

(1) 研究对象。本研究的研究对象为通过上海某教育自媒体招募的有兴趣参与数学学习力测量的1-9年级中小学生及其主要养育者。模型建构阶段的初步调研有75组家庭参与,模型实测阶段限于时间,截至本文数据分析成稿前,在平衡参与被试的人口统计学信息前提下,在参与了初步调研的75组家庭中随机抽选了20组家庭进行模型实测。

(2) 研究工具。本研究主要由测评模型构建与实测组成。模型主要通过数学应试能力测评卷设计、数学素养问题解决能力测评设计及非认知因素量表选择来进行构建,同时针对主要养育者设计了结构化访谈,以探索其期望对学生数学学习力影响的可能性。实测过程中会涉及问卷、情境实验、观察、访谈等综合研究方法。测量流程分为应试环节(做卷子)、情境实验环节以及非认知因素测评环节。应试环节本研究根据1-9年级沪教版数学上册部分进行了试卷出题,按照20分钟的题量

进行设计,该工具在正式使用前经上海体制内相应年级资深教师审核,并进行初步前测,能保证其难度适中具有代表性。情境实验环节采用自编数学问题解决情境测试活动,每2-3个年级共享一组活动,采用情境实验与观察法,对学生的行为表现进行记录与分析。非认知因素环节通过问卷量表和结构化访谈2种形式相结合达到测量目的,其中问卷量表参考^[6]数学学习经验调查问卷的数学焦虑分量表、^[7]改进编制的数学自我效能感量表、^[8]改进编制的数学创造力量表进行综合编制。

(3) 数据收集与处理。本研究采用微信及腾讯会议收集初步描述、访谈及过程观察等信息,并对其进行量化处理,量表部分通过问卷星收集。数据处理主要使用了Python编程,结合Numpy、Pandas、Matplotlib等数据分析及可视化包进行描述性统计及相关分析。

5 研究内容与发现

5.1 基于初步调研的描述性统计

在模型构建阶段的初步调研中,对于所有有意向参与后续研究的被试,要求主要养育者提供一段对于孩子数学学习情况的描述以及相关试卷、习题等材料(即数学学业方面详细信息)。对于提供了数学学业方面详细信息的75位学生,笔者进行了初步量化,初步量化涉及以下几个指标:

- ①孩子性别
- ②家长性别(父母哪一方先来咨询报名)
- ③年级(以2023年秋季开学计)

(1) 数学成绩初判:根据成绩好坏赋予1-5分,1分一档,数学成绩的判断来源于家长,包括家长口述、家长提供的测验成绩、排名等。5分代表成绩极好,3分代表中等成绩,1分代表成绩极差;完全不交代成绩情况的空白。

(2) 量化可靠度:根据量化来源可靠程度赋予1-5分——5分有提供成绩与大致班内排名情况;3分只提供了考试成绩;1分只有家长口头描述;完全不交代成绩情况的空白。对于量化后的数据,本研究进行了些统计分析 & 可视化,如下文所示。在递交信息的75组家庭中,本研究根据父母的性别和孩子的性别,做了二分,具体如下图:

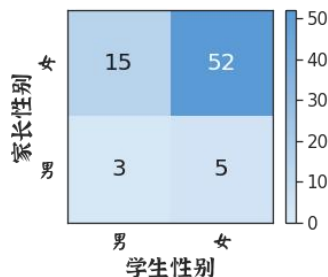


图1 学生与主要养育者性别的关系

孩子里面女生 57 人，男生 18 人；而家长中母亲 67 人，父亲 8 人。鉴于招募被试时有在介绍材料里提到数学焦虑的概念，并对于最终进入模型实测阶段的被试会有详细的针对性可操作建议反馈，从中做以下合理猜测，即：

①母亲比父亲更容易对子女的教育问题上产生焦虑

②女生在数学这门学科上更容易让家长焦虑。妈妈+女儿（75 组中占据 52 组）这个组合可以说是在咨询报名的对象中占据绝对主导地位了。

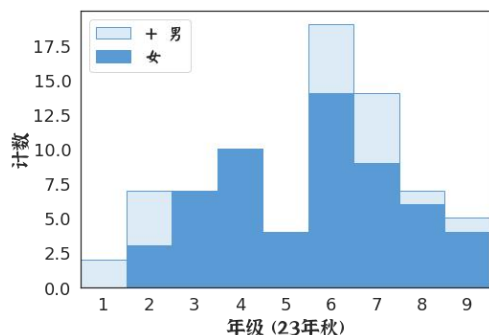


图2 初步调研参与者中的年级分布

如果分年级段来看的话，从上图中不难发现，初步调研阶段报名的六年级孩子是最多的。对于上海而言，小初阶段是5年+4年制的，也就是很多参与测评的主要养育者正在经历非常巨大的转变，根据实际情况推测可能是因为孩子刚上初中后跟不上数学课程节奏对数学学习力产生怀疑而焦虑的。然后这项研究在实际评测和沟通之前，笔者会在邀请父母给出关于孩子数学学习成绩的描述，并根据这些描述形成两个维度的数据：

①孩子数学成绩的分级，1-5

②父母描述孩子成绩的细节的细节可靠度，1-5

先来看关于数学成绩的数据分布：

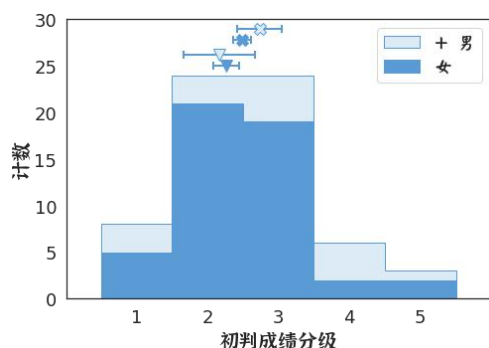


图3 初步调研参与者中的初判成绩分级分布

从上图成绩分布来看，绝大多数孩子的初判数学成绩分级都在2-3分档（略差-平均水平），笔者判断主要是因为不少家长没有直接交代清楚孩子的成绩状况（多为3分），或者只是口头上表示自家孩子数学成绩较差（多为2分）。需要说明的是，图中所有数据加起来不到总人数（75人）是因为有小部分

学生并没有该项数据。直接去计算平均值（上图中的叉标记）的话，可以看出男生的平均判定略高于女生，但考虑到误差的话也谈不上有什么差别，虽然符合刻板印象但也绝不像一般人所描述地那么极端。当然这里还是要说一下：这个样本本身就是筛选过的——毕竟是家长主动报名测评的，对孩子数学学习本身就不放心的家庭才会出现在这个统计样本中，这些结果是无法代表一般中小学男女生数学成绩分布的真实差别。接下去笔者把父母所描述的数学成绩的可靠度也加入进来一起做交叉分析，如下：

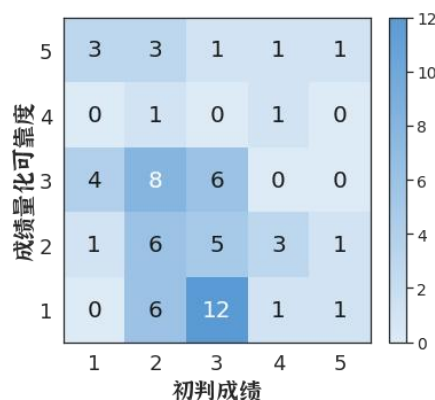


图4 成绩量化可靠度和初判成绩的关系

从上图笔者能看出，量化可靠度越低，成绩分布就越集中在中间值，而当父母描述的细节越丰富，成绩初盘的分级就开始逐渐整体左移。笔者在构建这项研究的时候，基于过往的日常课堂观察，形成了以下假设——学生的数学学习，是一个受三重因素影响的结果。这三重因素分别为：

①学生实际的数学力

②学生学习以外的相关状态

③家长的状态

因此笔者在积累被试报名的时候，就在有意识建立这些因素之间的可能关联，并搜集相关数据，以上只是一些例证。尽管以上的小规模数据现象的结果不应过度解读，但在此还是想作一个大胆地猜测：同样是对孩子数学学习存在焦虑的家长，如果是男孩子的话，往往是真正意义上的焦虑。针对这个猜测，后续的研究里也将逐步完善。

5.2 测评模型构建

在4.2中，笔者已提到研究工具构成及具体来源设计的情况，在此就流程及模型构建中的一些详细信息进行阐述。本测评包括三个部分：

（1）应试环节：20分钟完成一张由5-8道大题构成的测试卷（小学三年级及以下部分大题换成一系列数字计算题）。

（2）开放题面试环节：从2-3道非定向问题中选一道进行

简要地作答，给出解决思路即可。在此之后，将继续追问孩子剩下题目的解答思路。

(3) 非认知因素测评环节：通过量表与结构化面谈相结合的形式，综合了解孩子们不同方面焦虑、自我效能感以及数学兴趣。

所有环节进行完后每个被试将都会收到一份测评报告(下图为参考例)：



图5 测评报告例

其中，整体评价中会涉及到实际成绩、数学素养、实际畏难、孩子焦虑、数学兴趣五个方面。

(4) 实际成绩：依然5分制，但和初判成绩不同的是，此处0.5分一档(毕竟考察得更细致了)。最低1分，最高5分，分数越高越好。其中3分为平均水平。

(5) 数学素养：反映的是孩子的数学思维能力以及相关的理科素养。5分制，0.5分一档，3分为平均水平。这个维度的设置是用来判断孩子的数学潜力。

(6) 实际畏难：反映孩子实际的畏难程度。依然5分制，但和靠家长给出的畏难判定不同的是，此处0.5分一档，3分为平均，分数越低越好。

(7) 孩子焦虑：反映孩子实际的焦虑程度。5分制，0.5分一档，3分为平均，分数越低越好。

(8) 数学兴趣：反映孩子对数学底层原理、数学思维等方面的兴趣。5分制，0.5分一档，3分为平均水平。

5.3 实际一对一执行数学学习测评中的发现(数据统计层面，非个案)

先从现场做卷子所形成的实际成绩与之前初判成绩的对比讲起：

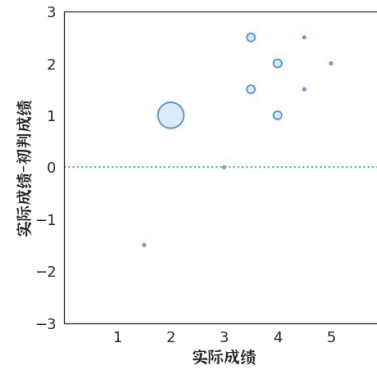


图6 实际成绩与之前初判成绩的对比

上图纵坐标是实际成绩减去上面一段中笔者根据主要养育者描述所给的初判成绩，所形成的差值。虚线为0，代表两者一致，即主要养育者的描述和实际情况相符。

如图所示，大部分的点都位于虚线上方的正值范围内。也就是说大部分的孩子实际成绩都高于基于主要养育者描述的预判成绩，由此可猜测，家长往往都低估了自己孩子的数学能力，幅度集中在1-2档。圆圈的大小代表有多少样本处于这个位置。如图，最集中的是实际成绩在中下水平(2档)，而家长的认知差值在1档(即家长认为孩子的数学水平是2-1=1档)。横坐标是实际成绩的绝对分值，左低右高。会发现孩子实际成绩越好，家长的低估幅度越大，两者正相关极其显著。第二部分的统计观察，是在数学兴趣、实际成绩和数学素养之间的两两关联。数学素养是在一张卷子以外，笔者和孩子进行的一系列关于数学的开放性问答。而数学兴趣则是笔者基于和孩子的整体沟通下来，再结合量表根据观感形成的量化指标。整体而言，这三个因素两两之间都有正相关性。

然而三者之间两两正相关的程度有区别。先来数学兴趣和实际成绩：

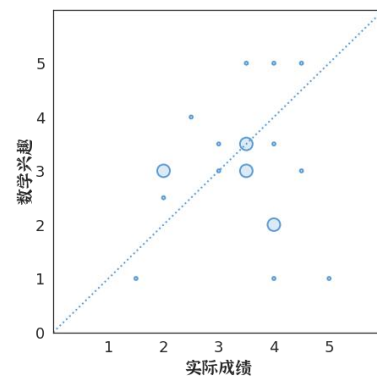


图7 实际成绩与数学兴趣的关系

有正向相关，但比较弱，看图也看得出有一些随机性。

数学素养和实际成绩之间也呈现类似的情况：

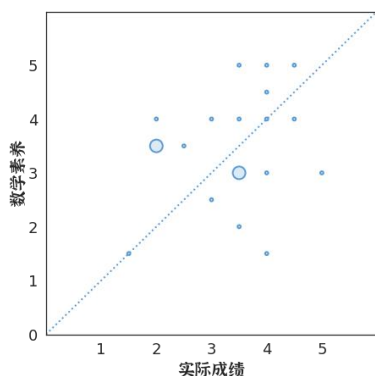


图8 实际成绩与数学素养的关系

实际关联度最高的,出现在数学兴趣和数学素养之间,而且本组因素关联之间最集中的人群,也出现在本图中:高兴趣高素养。

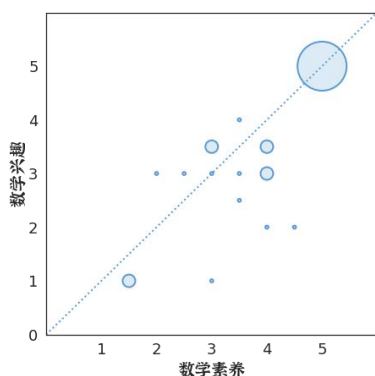


图9 数学兴趣与数学素养的关系

5.4 研究讨论与建议

5.2中几张图的相关现象,推测也许与笔者对孩子的数学素养的考察以及数学兴趣判定方式有关。小初阶段的数学更加偏重算术本身,以中考为导向的数学训练可以靠重复性刷题来获得较好的应试成绩。而笔者的面试环节倾向于考察孩子将数学融入到日常思维的能力,以及孩子在已有的数学知识基础上,用演绎思维推导理论的潜力,而这些并不能体现在小初阶段的考试上,但会影响到后续高中阶段。数学兴趣的判定上,

参考文献:

- [1] Schoenfeld,A.H.(1985).Mathematical problem solving.Academic Press.
- [2] Boaler,J.(2006).How a detracked mathematics approach promoted respect,responsibility,and high achievement.Theory Into Practice,45(1),40-46.
- [3] Embretson,S.E.,&Reise,S.P.(2000).Item response theory for psychologists.Lawrence Erlbaum Associates.
- [4] Pellegrino,J.W.,&Hilton,M.L.(Eds.).(2012).Education for life and work:Developing transferable knowledge and skills in the 21st century.National Academies Press.
- [5] Gee,J.P.(2003).What video games have to teach us about learning and literacy.Palgrave Macmillan.
- [6] 吴明隆.(2000).SPSS 统计应用实务.中国铁道出版社.
- [7] 潘佳瑶.(2020).农村地区高中生数学焦虑、数学自我效能感与数学成绩关系的研究硕士学位论文,西南大学).
- [8] 韦胜男.(2023).初中生数学焦虑、数学自我效能感与数学创造力的关系研究硕士学位论文,湖北大学).

笔者重点区分了孩子所谓的数学兴趣,究竟是对数学底层原理本身感兴趣,还是单纯因为成绩上的正反馈给自己的满足感让自己误以为对数学感兴趣。而由于在整个研究过程中接触孩子的主要养育者比较多,在此也简要说几点从研究出发给到主要养育者的建议:

(1) 去相信孩子的潜力,眼前的“鸡飞狗跳”不等于没有未来,平常心面对孩子的数学学习,上面的数据大家也看到了,会寻求的帮助的各位大多孩子的数学学习问题并没有想象中的大,有时家长的过度担忧才是未来的“炸弹”。

(2) 小学阶段,尤其小学 1-3 年级的数学学习问题,通常数学知识无关,基本上是生理、心理方面等非认知因素问题。

(3) 无论体制内外,去尊重一般意义上的教育规律总是没错的,这里不是指学校说什么就是什么,而是有时大家觉得甚至有些务虚的部分,比如说,每个孩子都有自己成长的节奏,可以“养花”别“驯兽”,最终这一切会反哺到数学学习上的。

5.5 研究局限性与展望

限于时间,本研究虽说有每一个个案的详细材料,但目前并未就个案做单纯的个案分析研究,后续会在时间充裕的情况下,对已收集案例详细数据进行个性化的个案研究。目前测评样本量依旧较小,后续也会继续收集被试样本数据,并不断优化现有模型。目前测评模型开发至 1-9 年级,后续如有条件将扩展至高中阶段,看到测评过程与结果反馈真的可以让一个孩子乃至一个家庭变好,感觉的确也是很光明的项目了。

6 研究结论

本文研究结论如下:研究建立了一个青少年数学学习力模型,该模型有助于全面理解数学学习过程,并为学生、家长、教师和政策制定者提供参考和指导,以促进每个学生的个性化发展。研究发现,数学兴趣和数学素养之间存在正向关联,且这两者与实际成绩也呈正向关系,但关联度较弱,并存在一定的随机性。