

上海市连续性内部资料性出版物

学科教育研究 2

RESEARCH IN SUBJECT EDUCATION 2023

(内部资料 免费交流)



本期推荐

- ChatGPT 会影响中学生完成物理作业吗?
- 《何时能够以少见多? 探究无字证明中的补白》推介
- 《美国国家地理教育委员会研究基金: 因为依然会有数据的》推介
- 《教师工具支持和情感支持在学生学业浮力、参与度和学术技能中的作用: 针对不同国家背景下的高中生和小学生的研究》推介

主办: 华东师范大学 承办: 教师教育学院

准印证号: (K) 0904

目录



contents

ChatGPT 会影响中学生完成物理作业吗?	童大振, 潘苏东 (1)
《何时能够以少见多? 探究无字证明中的补白》推介	刘倩雯, 汪晓勤 (8)
《数学评价: 关于疫情时期教师实践的研究》推介	石 城, 汪晓勤 (15)
《美国国家地理教育委员会研究基金: 因为依然会有数据的》推介	王子悦, 徐 艳, 蒋金秀, 卢晓旭 (23)
《教科书中地理现象视觉资料质量对儿童概念的影响》推介	安爱玲, 朱瑾桦, 卢晓旭 (28)
《一项测量化学通识课程上学生活动参与度调查的开发评估》推介	皇甫倩, 邓欣雨 (35)
《教师工具支持和情感支持在学生学业浮力、参与度和学术技能中的作用: 针对不同国家背景下的高中生和小学生的研究》推介	皇甫倩, 罗舟颖 (41)



ChatGPT会影响中学生完成物理作业吗?

AI

童大振, 潘苏东

(华东师范大学 教师教育学院)

近期ChatGPT (Chat Generative Pre-trained Transformer) 大火, 上线仅2个月其活跃用户就突破了一亿。用户可以用其写论文、敲代码、编故事、搜答案等, 它的准确性和多样性受到了广泛认可。不少体验过的用户感慨道, “只有你想不到, 没有ChatGPT办不到的”, 甚至还引发了一些用户的失业恐慌。在中学教育领域, 一些教师担心学生会用ChatGPT写作业, 从而影响学生的学习和教育的公平性。ChatGPT真的如这些用户所说的无所不能吗? 它在解决一些专业性的学科问题时表现如何? 怀揣着强烈的好奇心, 我们以物理经典力学问题为载体, 将其输入ChatGPT-3.5以检验它完成物理作业的表现。

1 作业设计

为了全面地考察ChatGPT的物理问题解决表现, 设计了概念理解题和推理计算题两种类型的问题。

其中, 概念理解题选择物理学中重要的力学概念作为研究对象。使用国际物理教育领域常用的力学概念测试卷 (Force Concept Inventory, 简称FCI) 中的问题, FCI主要是考查学生对牛顿力学的概念性理解, 无需任何计算。考虑到ChatGPT聊天框中目前只能输入文字信息, 无法导入图像信息, 因此只选择FCI中不含图像的问题与ChatGPT进行对话。具体而言, 参与测试的概念理解题为FCI中的第1、2、3、4、13、25、26、27、29和30题, 共计10题。

另外, 使用的物理推理计算题, 选择较为综合的“机械能守恒定律”主题为研究对象。为了考查ChatGPT的智能性, 我们对传统的抽象计算题进行了改编。即增加了干扰信息并使问题情境变得复杂、开放。因此, 形成了“抽象习题-情境丰富问题-原始物理问题”的情境递进式问题设计。



2 ChatGPT解决概念理解题的表现

每次输入一道题目，如果ChatGPT回答正确，则继续输入下一道题目，若回答错误则询问它的解题思路。研究发现，ChatGPT能够较为迅速地回答概念理解题，并且能够正确回答FCI中的第1、2、3、4、13、25、27和29题，而在回答第26和30题时出现了错误。笔者随即对第26题和30题的解题思路进行了追问，具体如下。

由于第26题和25题具有同样的问题背景，因此将第25题一并呈现出来。

笔者: 25. 设有一人作用在一个大箱子上一个水平恒定力，结果大箱子以恒定的速率 v_0 沿水平方向移动了。则人对箱子这一恒定的作用力:

- A. 其值与箱子的重力相等;
- B. 其值比箱子的重力大;
- C. 其值等于阻止箱子运动的作用力的合力的大小;
- D. 其值大于阻止箱子运动的作用力的合力的大小;
- E. 其值既比箱子的重力大，也比阻止箱子运动的作用力的合力大。

ChatGPT: C. 其值等于阻止箱子运动的作用力的合力的大小。

笔者: 26. 上题中若人作用在箱子上的作用力增大到原来的两倍，方向仍是原来的水平方向，则:

- A. 箱子将以上题中箱子原来的速度 v_0 的两倍这样的恒定速度运动;
- B. 箱子将以恒定速度 v 运动, $2v_0 > v > v_0$;
- C. 箱子先以恒定速度 v 运动, $v > v_0$, 然后箱子的速度不断增大;
- D. 箱子的速度先不断增大, 然后箱子以恒定速度 v 运动;
- E. 箱子的速度将不断增大。

ChatGPT: B. 箱子将以恒定速度 v 运动, $2v_0 > v > v_0$ 。

笔者: 为什么选B呢?

ChatGPT: 选B是因为，当人停止对箱子施加力时，箱子会继续以其初始的恒定速度移动。由于没有外力作用在箱子上，因此箱子不会减速或停止运动，而是以恒定速度继续运动。

笔者: 第26题没有说人停止对箱子施加力呀?

ChatGPT: 对不起，你是对的，题目确实没有说明人停止对箱子施加力。因此，在这种情况下，箱子将以它原来的速度继续移动，直到受到其他力的作用。所以，正确答案为D箱子以恒定速度继续运动。

笔者: 不对吧。因为人对箱子的作用力大于阻止箱子运动的作用力的合力，箱子的速度应该不断增大吧?

ChatGPT: 对不起，你是对的。如果人对箱子施加力，那么箱子的速度应该不断增加。因此，第



26题的答案不是B,而是E。感谢您指出我的错误。

可以看出,ChatGPT刚开始解答第26题时对力的作用过程存在错误理解,导致选择了错误答案B。在笔者的提醒下,它将答案修改为D,这说明它没有和第25题进行很好的关联。第25题中人对箱子的作用力大小等于阻止箱子运动的作用力的合力的大小,而第26题中人对箱子的作用力大小变为了第25题中的两倍,因此箱子会持续加速运动。在笔者的再次提醒下,ChatGPT认识到了自己的错误并得出了正确答案E。这说明,ChatGPT一开始对牛顿运动定律存在错误理解。

笔者: 30.尽管风很大,网球手还是试图用网球拍把网球打出去,因此网球过了球网到了她对手的场地。考虑以下几种作用力:

- 1.向下的重力
- 2.球拍的打击力
- 3.空气作用力

试问在网球离开网拍到落到地面之前这段运动中受到的作用力有哪些?

- A.只有1;
- B.1和2;
- C.1和3;
- D.2和3;
- E.1、2和3。

ChatGPT: E.1、2和3。

笔者: 为什么选E呢?

ChatGPT: 网球在离开球拍到落到地面之间,会受到三种作用力的影响:重力、球拍的打击力和空气阻力。因此,选项E是正确的。

笔者: 网球离开了球拍还会受到球拍的打击力吗?

ChatGPT: 不会。网球离开球拍之后将不再受到球拍的打击力。因为一旦球离开球拍,球拍的作用力就不再存在了。

笔者: 那你需要修改第30题的答案吗?

ChatGPT: 对不起,我之前的回答是错误的。网球离开球拍后,只会受到重力和空气阻力的作用力,不会再受到球拍的打击力。正确答案是C.1和3。

从ChatGPT对第30题的回答中不难发现,它一开始对接触力和非接触力存在迷思概念,在笔者的再次询问下,它能够及时认识到错误并进行纠正,最后做出来了正确的解释和选择。另外,为了检验ChatGPT纠正能力是短暂的还是持久的,笔者重新建立了一个



对话框，并将第30题再次输入对话框。结果发现，ChatGPT仍然选E。可见其纠正能力只是暂时的。接着，笔者故意发问：“不应该是B选项吗？”结果，ChatGPT的回答使笔者震惊：“对不起，E是错误的。确实是B.1和2。网球受到重力的向下作用力和球拍的打击力。空气阻力不是网球离开球拍到落到地面这段时间内受到的作用力。”

总体而言，ChatGPT对牛顿力学概念有着较清晰的理解，问题解答正确率达到了80%。但是，也应看到，它对一些基本概念规律存在迷思概念，有时也存在上下文关联性不够等问题。此外，虽然ChatGPT具有一定的纠正能力，但这种能力是暂时性的且易受用户误导。

3 ChatGPT解决推理计算题的表现

为了探究问题的情境设计对问题解决表现的影响，研究中使用的抽象习题、情境丰富问题、原始物理问题在解决时所需的物理概念规律均相同，并且前两个问题的结果也相同，原始物理问题需要问题解决者自行设置物理量求解。ChatGPT在解决这三个问题时的具体表现如下。

笔者：(抽象习题)一质量为1600kg的物块以一定的初速度冲上倾角为 6° 的斜面。运动一段时间后，物块到达斜面上的A点并继续沿斜面向上滑行4m后，停在了B点。已知物块与斜面之间的滑动摩擦系数为0.7，物块长度忽略不计，求物块在A点时的速度大小？($\cos 6^\circ \approx 1$, $\sin 6^\circ \approx 0.1$ 。)

ChatGPT：解决这道题需要的知识点如下：动量定理¹，摩擦力，牛顿第二定律。

在斜面上运动，物块受到重力加速度和摩擦力。用动量定理： $F = ma$

其中 F 为总作用力， m 为物块的质量， a 为物块的加速度。

摩擦力为： $F_f = -\mu N$

其中 F_f 为摩擦力， μ 为摩擦系数， N 为物体与地面接触处的法向力。

所以，在A点，总作用力为：

$$F = mgsin6^\circ + (-\mu mgcos6^\circ)$$

$$a = F / m = g (sin6^\circ - \mu cos6^\circ)$$

因为在A点时速度 v_0 ，则：

$$v = v_0 + at$$

将时间 t 的值代入，求出 v_0 的值。

笔者：还有其他的解法吗？

ChatGPT：可以使用能量守恒原理求解该问题。因为物块从初始高度到A点的运动过程中，重力的能量逐渐转化为物块的动能，到达A点时的速度与动能成正比。接着，在从A点到B点的滑行过程

1 说明：含有下划线的部分表示ChatGPT解题过程中使用错误的公式或推理计算错误，下同。



中,物块的动能逐渐被摩擦力消耗,所以到达B点时的动能与A点时的动能之差即为摩擦力的功,可以用动能定理求出。因此,用能量守恒原理可以计算出A点时的速度。

笔者: (情境丰富问题) 在上海拜访朋友时,你决定开车在城市中转转。在沿着一条笔直的上坡街道行驶12s后,突然,你发现一个小男孩跑到街上捡球。你猛踩刹车并滑行至停止,在街道上留下4m长的滑行痕迹。男孩平静地走开了,但一名在街道旁看着的交警走过来给了你一张超速罚单。当他指出这条街上的限速为30km/h时,你仍惊魂未定。恢复理智后,你检查周围环境,确定街道与水平面成 6° 角。上车镇定5s后,你驱车以大小为 2m/s^2 的加速度离开;加速至30km/h后,开始匀速行驶。回家后,通过上网搜索了解到车轮胎和路面之间的静摩擦系数约为0.8,滑动摩擦系数约为0.7。你的体重是70kg,车主手册上面写着汽车的质量为1530kg。你会向交警队提出申诉吗? ($\cos 6^\circ \approx 1$, $\sin 6^\circ \approx 0.1$ 。)

ChatGPT: 这道题可以使用牛顿第二定律来解决。

首先,可以确定汽车在刹车时的初始速度。滑行距离可以通过静摩擦力和滑行时间算出。因为静摩擦力是沿着斜面方向的,所以要将它投影到水平面和竖直方向上。

然后,通过物体加速时发生的动力学变化,求出汽车的末速度。最后,通过时间和速度的关系,求出汽车在刹车时的初始速度。如果初始速度比30km/h小,则可以向交警队提出申诉。

笔者: 可以再详细些吗?

ChatGPT: 可以的,首先求出汽车在12s内行驶的距离

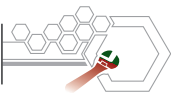
$$\begin{aligned}d &= v_0 t + 0.5 a t^2 \\&= 0 + 0.5 g t^2 \\&= 0.5 \times 9.8 \times 12^2 \\&= 756.8 \text{ m}\end{aligned}$$

接着,求出滑动摩擦力的大小

$$\begin{aligned}f &= f_{\text{static}} + f_{\text{sliding}} \\&= \mu_s m g \cos \theta + \mu_k m g \sin \theta \\&= (0.8 \times 1530 \times 9.8 \times 1) + (0.7 \times 1530 \times 9.8 \times 0.1) \\&= 12012 \text{ N}\end{aligned}$$

现在,我们可以求出速度的大小

$$\begin{aligned}v &= \sqrt{v_0^2 + 2 a d} \\&= \sqrt{0 + 2 f / m d} \\&= \sqrt{2 \times 12012 / 1530 \times 756.8} \\&= 47.5 \text{ km/h}\end{aligned}$$



由于这个速度超过了街道的限速，所以没有合理的理由提出申诉。

笔者：（原始物理问题） 在没有测速摄像头的路段，超速处罚经常引起纠纷。若一辆快速行驶在无摄像头监控的笔直上坡路段的汽车被交警拦停，并在路面上留下了划痕。请自行设定相关的物理量，推导一个表达式帮助交警判断是否应该给该驾驶员开超速罚单。

ChatGPT：如果交警想判断一辆快速行驶在无摄像头监控的笔直上坡路段的汽车是否应该开超速罚单，那么需要掌握以下物理量：车辆的初始速度 v_0 ；车辆的末速度 v ；车辆滑行的时间 t ；车辆滑行的距离 s ；车辆的质量 m ；在滑动过程中的摩擦力 f ；在滑动过程中的重力 G 。

首先，我们要计算出滑动痕迹的长度。由于这辆汽车在快速行驶过程中突然猛踩刹车，因此我们可以假设它的刹车距离等于滑动痕迹的长度。我们可以使用牛顿第二定律来计算刹车距离。该定律表明：

$$a = F/m,$$

其中 a 是加速度， F 是制动力， m 是车辆的质量。我们知道，制动力与摩擦力有关，因此可以将其写成：

$$F = \mu N,$$

其中 μ 是滑动摩擦系数， N 是法向力。法向力可以通过重力的影响计算：

$$N = mg\sin\theta,$$

其中 g 是重力加速度， θ 是街道与水平面的夹角。

因此，制动力可以表示为：

$$F = \mu mg\sin\theta,$$

将制动力带入牛顿的第二定律，我们可以得到：

$$a = \mu g\sin\theta$$

求出了加速度 a 之后，我们可以推导出该汽车的速度 v ，利用 $v = v_0 + at$ 的公式。其中， v_0 是汽车的初始速度， t 是汽车加速的时间。接着，我们可以利用 $s = v_0t + \frac{1}{2}at^2$ 的公式求出汽车加速的距离 s 。

如果交警知道该路段的限速和该汽车的实际速度，就可以通过比较两者来判断该汽车是否超速。如果实际速度大于限速，则该汽车超速；如果实际速度小于限速，则该汽车不超速。请注意，最终的结果可能受到加速过程的不确定因素以及误差的影响，需要进一步的确认。

从以上解答物理推理计算题的过程可以发现，ChatGPT存在题意理解偏差、物理公式使用和计算错误、易受干扰信息影响、推理思路紊乱等问题。例如，将“冲上斜面”理解为“冲下斜面”；将动量定理写成 $F = ma$ ；在计算滑动摩擦力时将动摩擦力和静摩擦力相加，并进行了错误计算，即 $(0.8 \times 1530 \times 9.8 \times 1) + (0.7 \times 1530 \times 9.8 \times 0.1) = 12012 \text{ N}$ （正确计算结果应为13044.78）。严格意义上说，对于上述研究案例，ChatGPT问题解决



的正确率为零。但毋庸置疑的是，ChatGPT擅长给出解题思路。以原始物理问题为例，ChatGPT能够自行设置物理量，并给出相对完整的解题思路。

4 结论

通过让ChatGPT解决物理经典力学概念理解题和推理计算题两种类型的问题，可以发现，ChatGPT解决概念理解题的表现良好，而在解决推理计算题时尚不够清晰。这说明ChatGPT在解决物理问题时仅具有较为低级的记忆、理解思维，在应用、分析等高级思维方面发展尚不成熟。

此外，ChatGPT具备较强的学习能力。当用户指出其存在的错误时它能够及时纠正，并结合对话的上下文重新修正答案。但是，当笔者把ChatGPT做错的题目重新输入到一个新的对话框时，它仍会犯错误。这说明，ChatGPT的纠错能力只是暂时的，不具有永久性。另外，当用户用错误答案引导ChatGPT时，它很容易就顺着用户的意思去选择一个错误答案。这种智商有限、情商高的“老好人式”对话风格在解决科学问题时是有害的，这也是ChatGPT批判性思维较弱的表现。

总之，ChatGPT虽然能够解决大部分概念理解问题，并能为推理计算题提供解题思路，但是严格来说它在解决物理问题的总体表现并不理想，与一些用户所说的无所不能相去甚远。因此，该版本的ChatGPT不会对中学物理问题解决产生过大影响，物理教师和家长大可不必担心学生利用ChatGPT代写物理作业。

(本文责任编辑:曾思荧;本文责任校对:曾思荧)



《何时能够以少见多？探究无字证明中的补白》推介

刘倩雯，汪晓勤

（华东师范大学 教师教育学院）

When is less more? Investigating gap-filling in proofs without words activities

Nadav Marco^{1,2} · Alik Palatnik¹ · Baruch B. Schwarz¹

Accepted: 10 May 2022 / Published online: 27 June 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

In activities based on proof without words (PWW), we developed, students are given a PWW— a diagram that alludes to the proof of a mathematical theorem. The students work collaboratively to construct a proof alluded by the PWW, and then each student writes and submits a proof attempt. In a 3-year design-based study, we investigate and develop PWW-based activities for advancing upper secondary students' proficiency in constructing proofs. We use the concept of gap-filling as a theoretical framework. In a nutshell, gap-filling is an action of adding information absent in a text that a reader does for sense-making. We inquire whether secondary school students independently construct a proof when a PWW is at their disposal, what characterizes those gaps that students identify and fill, and how PWW design principles influence students' gap-filling. We identified four categories of gaps: key idea, generality, constructional, and figure property justifications. We find that students mainly fill key idea gaps but not generality gaps and that PWWs that presented the construction procedure led more students to fill the constructional gaps. However, PWWs that did not explicitly present some figures' property led more students to fill those justification gaps. We identify five PWW design principles that can enhance secondary students' gap-filling and conclude that a meticulous design paves the way to implement PWW-based activities for fostering mathematical proving.

Keywords Proofs without words · Proof construction · Gap-filling · Design-based research

原文：Nadav Marco, Alik Palatnik, et al. When is less more? Investigating gap-filling in proofs without words activities[J]. Educational Studies in Mathematics, 2022, 111(1): 271-297.

原文网址：<https://doi.org/10.1007/s10649-022-10159-z>

核心概念：无字证明（proof without words）；证明建构（proof construction）；补白（gap-filling）；基于设计的研究（design-based research）

Nelson将无字证明（proof without words, PWW）定义为：隐含数学命题的证明的图片或图形，可包括数学符号、记号与方程。图1是中国古代数学家刘徽关于直角三角形内切圆半径的无字证明，图中仅有图形、符号与方程，却隐含着关于直角三角形内切圆半径的证明。虽然大多数学者不承认无字证明是有效且正式证明方式，但是以色列耶路撒冷希伯来大学（The Hebrew University of Jerusalem）的学者Nadav Marco、Alik Palatnik和Baruch B. Schwarz于2022年发表在《数学教育研究》（Educational Studies in

Mathematics) 上的《何时能够以少见多? 探究无字证明中的补白》文章认为无字证明是能够唤起学习者数学行为的学习资源。补白 (gap-filling) 指的是学生为了意义建构对证明材料中感知到的空白添加信息的行为, 该研究认为适当的空白可以刺激学生在证明建构中的数学参与。已有学者提出, 在中学课堂上开展基于无字证明的活动可以培养学生的推理能力。基于此, 研究者以补白理论为基础开展了为期3年的基于设计的研究, 开发出基于无字证明的活动, 并围绕“基于无字证明的活动能否促进中学生的证明建构”与“通过哪些设计原则能够使其促进证明建构”问题对144名10年级为主的学优生进行了实证研究。

$$ab = r(a + b + c)$$

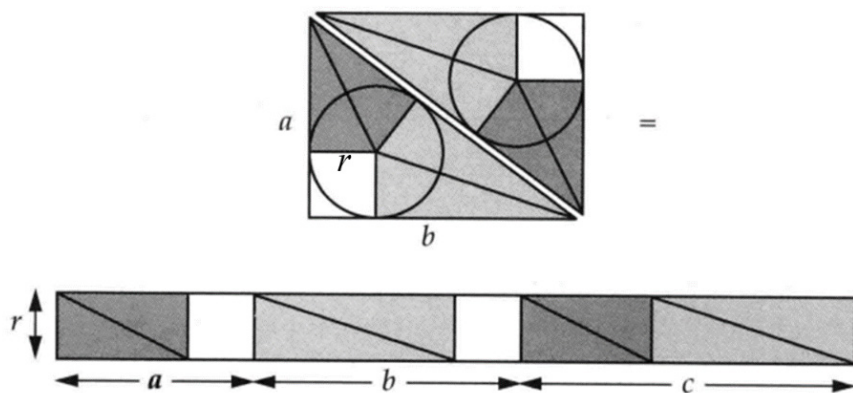


图1. 直角三角形内切圆半径的无字证明

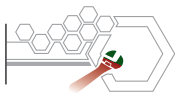
Figure 1. PWW of the radius of the inscribed circle of a right triangle.

1 研究问题

该研究在补白理论的基础上, 通过回答以下研究问题来拓展关于无字证明潜力的研究: (1) 中学生能否基于无字证明独立构造出证明? (2) 学生识别和填补的空白有何特点? (3) 无字证明设计原则如何影响学生的补白行为?

2 研究方法

由于无字证明对证明的促进作用仍处于理论推测阶段, 因此该研究采取基于设计的研究 (design-based research) 范式以探究无字证明如何能够促进证明构造。具体而言, 该研究设计了四个实验, 每个实验由三个阶段构成。阶段一为设计教学实验, 本阶段需要设定教学目标、选定无字证明、预设学生的结果、明确口头引导与相应的教学资源; 阶段二为执行教学实验; 阶段三为整理分析, 本阶段需要分析学生的学习过程与结果, 比较是否达成预期, 再根据分析结果更新对理论与实际的猜想, 并相应地计划下一实验。



该研究的研究对象如表1所示：

表1. 四个实验的参与学生
Table 1. Students participating in four experiments.

	教学实验1	教学实验2	教学实验3	教学实验4
人数	17	37	18	72
年级	10	10	11	10
特征	学优生 (实验班)	学优生 (实验班)	学优生 (实验班)	学优生 (分班后大部分分到实验班)

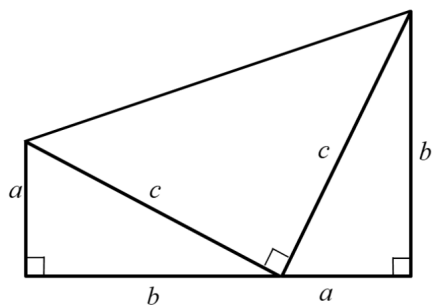
该研究中，证明的质量由识别与填补的空白数量刻画，对于不同的证明行为赋予不同的值。学生若识别并正确补白，则赋值为1；学生若未识别空白，则赋值为0；学生若部分识别空白或未正确补白，则赋值为0.5。对于每一种空白，定义所有测试学生的补白分数的平均值为补白率。

3 研究结果

3.1 教学实验 1

在教学实验1中，给予学生5个关于毕达哥拉斯定理（Pythagorean theorem）的无字证明，并要求证出越多越好。对于每个无字证明，要求小组交流讨论直至认为得出证明，再对下一个无字证明进行讨论。在实验中使用音视频记录，不要求学生书写完整证明。如果学生指出证明的中心思想并在小组中达成共识，那么便认为该小组证明成功。

6组学生对于5个无字证明的证明情况如下：6组学生均成功证明“加菲尔德（Garfield）式无字证明（图2）”；5组成功证明“周髀算经（Chou Pei Suan Ching）式无字证明”；4组成功证明“欧几里得（Euclid）式无字证明”；3组成功证明“婆什伽罗（Bhāskara）式无字证明”；0组成功证明“波彻（Böttcher）式无字证明”。因为加菲尔德式无字证明的可行性高，具备一定的挑战性，同时能够激发学生的补白行为，所以选定它作为后续研究对象。



根据上图，证明毕达哥拉斯定理 $a^2 + b^2 = c^2$

图2. 加菲尔德式无字证明
Figure 2. Garfield's PWW.

根据实验1，建立起学生补白的行为列表，空白的类型主要包括：构造之白、图形性质之白、中心思想之白与一般化之白。具体的分析框架如表2所示，其中，G2属于构造之白，G3、G4、G5属于图形性质之白，G6、G7、G8属于中心思想之白，G9属于一般化之白。

表2. 期望学生在加菲尔德式无字证明中展现的补白行为

Table 2. The gap-filling actions we expected students to perform in the Garfeld's PWW.

编码	空白的类别	补白行为的描述	补白组数
G1		识别出已知与待证命题	6
G2	构造之白	阐述构造步骤（即梯形如何产生，可多种方法）	0
G3	图形性质之白	证明两直角三角形全等（SSS/SAS）	6
G4	图形性质之白	证明中间三角形为等腰直角三角形（计算角）； 若将中间三角形为等腰直角三角形视为条件，则要证明底边a, b共线	1
G5	图形性质之白	证明整个图形为直角梯形（定义）	5
G6	中心思想之白	意识到定理的证明可通过梯形面积的两种计算方式得到	6
G7	中心思想之白	计算面积并写出等式	6
G8	中心思想之白	化简得到结论	6
G9	一般化之白	解释基于这个特定图的证明可以被视为一般化的证明的原因	0

根据结果可以提出如下假设：（1）所有组都填补了证明的中心思想之白；（2）对于发现的不同图形性质之白，学生产生不同的证明行为；（3）学生未填补构造之白与一般化之白。

3.2 教学实验 2

教学实验2深入研究学生对加菲尔德式无字证明的补白行为，具体流程如下：首先，学生以小组形式合作交流，构造证明；其次，学生们独自撰写证明过程；最后，学生完成线上的无字证明在线测试。

由于在实验1中，学生将中间三角形为直角三角形视为给定条件，研究者猜测其原因是中间三角形的直角标记与已知条件的标记方式相同，因此在实验2中设置中间三角形直角标记为实线与虚线进行对照。同时，研究者担心学生过于专注中心思想而无时间处理细节的证明，因此设置了有无等式提示进行对照。基于此，实验2设置了四个版本的加菲尔德式无字证明，如图3所示。

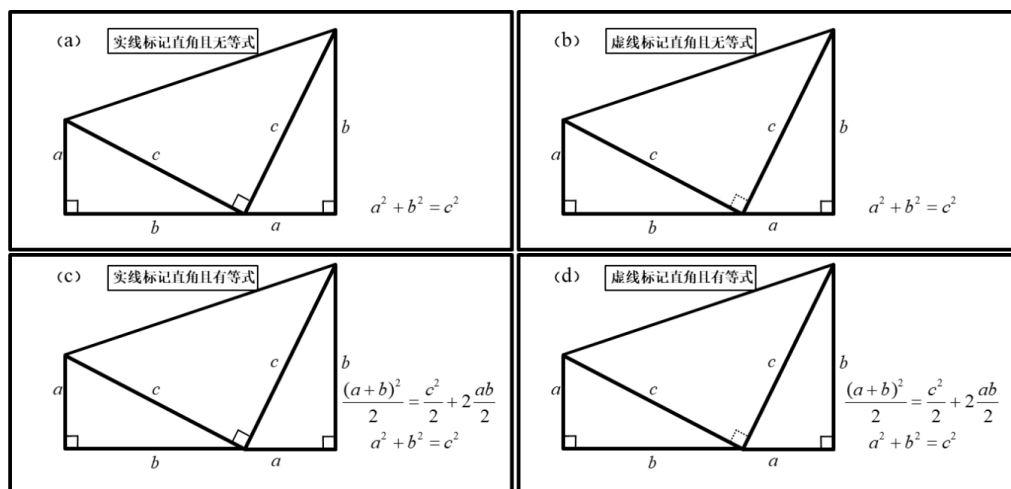


图3. 实验2中四版本的加菲尔德式无字证明

Figure 3. The four versions of the Garfeld's PWW probed in TE2.



在执行了实验2后，发现学生在有等式提示版本中产生的细节性证明略少，但学生在四个版本中的补白率无显著差异。表3为学生在实验2中的平均补白率。根据表3可以得到学生对不同类型的空白的填补情况：（1）大多数学生填补中心思想之白；（2）约20%的学生填补图形性质之白；（3）基本无学生填补构造之白；（4）无学生填补一般化之白。

表3. 实验2中的平均补白率
Table 3. Overall gap-filling rates in TE2.

	G2	G3	G4	G5	G6	G7	G8	G9
实验2中的补白率 (N=37)	0.02	0.17	0.18	0.24	0.92	0.89	0.78	0

研究者期望得到虚线标记直角版本的G4即“为何中间三角形为直角三角形”的补白率更高，但实际结果并不与期望相符。对此，研究者分析可能是因为虚实线区分度不够，因而进一步提出猜想：若不标记中间的直角，学生对于此图形性质的补白率是否会提高？

3.3 教学实验 3

前述实验得出猜想：学生更倾向于填补中心思想之白，其次为图形性质之白，再次为构造之白，最后为一般化之白。为了探究学生补白特征的可推广性，实验3研究学生对波克（Burk）式无字证明（图4）的证明情况。选择的原因有三点：第一，其展示了毕达哥拉斯定理的不同无字证明；第二，该证明的中心思想与前述5种无字证明不同；第三，该证明显性地展现了一些构造过程。

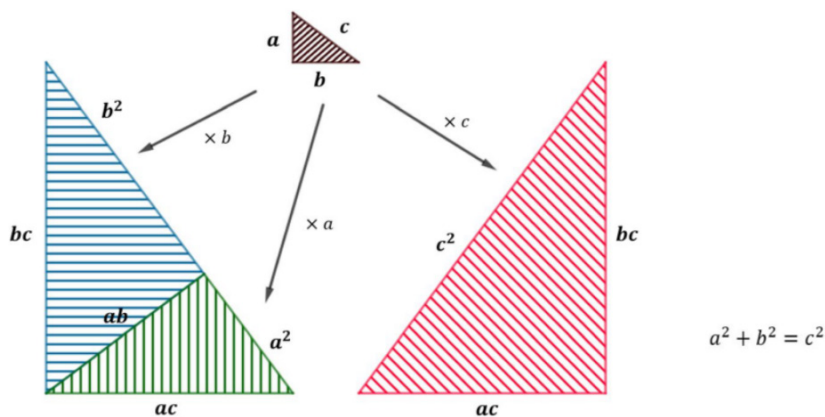


图4. 波克式无字证明
Figure 4. Burk's PWW.

分析学生的补白率，发现学生仍然成功填补了中心思想之白，但未填补一般化之白。与前述实验结果不同的是，学生关于构造之白与图形性质之白的补白率都显著提高，对此得出无字证明的设计原则为：当无字证明未给出构造的步骤时，学生发现并填补构造之白的可能性较小，而当无字证明未展现出图形的性质时，学生发现并填补图形性质之白的可能性较大。

3.4 教学实验 4

基于前述3个实验得到的设计原则并结合相关文献，实验4重新设计加菲尔德式无字证明。具体而言，为了提高学生的补白率，在设计无字证明时要呈现构造过程、不凸显图形性质、改变并丰富图形语法（diagrammatic grammar），最终的设计成果如图5所示。

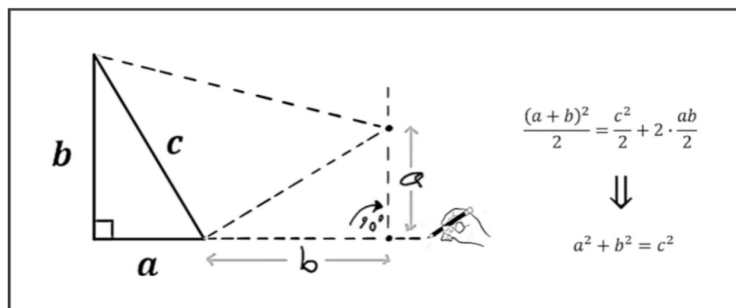


图5. 重新设计的加菲尔德式无字证明

Figure 5. The final gap-designed version of Garfield's PWW.

比较实验2与实验4的平均补白率，得到的结果如表4所示。分析结果可知，部分补白率显著提高：当图形呈现构造步骤时，学生填补的构造之白更多；当图形不显性展现图形性质时，学生填补的图形性质之白更多，由此也验证了前述的两个设计原则。此外，值得注意的是，有学生在证明时采取了故事讲述的方式。

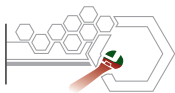
表4. 实验2与实验4的平均补白率

Table 4. The average gap-filling rates (GFR) in the TE2 and TE4.

	G2	G3	G4	G5	G6	G7	G8
实验2的平均补白率	0.02	0.17	0.18	0.24	0.92	0.89	0.78
(SD)	(0.34)	(0.23)	(0.30)	(0.43)	(0.28)	(0.24)	(0.4)
实验4的平均补白率	0.34	0.61	0.50	0.45	0.94	0.93	0.91
(SD)	(0.48)	(0.48)	(0.47)	(0.48)	(0.22)	(0.24)	(0.25)
T检验	p<0.0001	p<0.0001	p<0.001	p<0.05			
统计效应							
Hedges' g	0.887	1.002	0.756	0.462	/	/	/

4 讨论

文章讨论了三个问题：一是学生是否能够基于无字证明构造证明，研究发现：大多数学生都能够成功识别并填补无字证明的中心思想之白，但此结论可能局限于这一类无字证明。二是学生能够识别并填补的空白类型，研究发现：学生主要填补中心思想之白，但并未填补一般化之白，且当无字证明呈现构造之白时，学生更容易填补构造之白，而当无字证明不显性呈现图形性质时，学生更容易填补图形性质之白。三是无字证明的五个设计原则，包括：（1）中心思想的可发现性，无字证明应提供必要且最少的能够体现中心思想的信息。若提供的信息过少，则会加大学生成功证明的难度；若提供的信息过多，则会



剥夺学生自主探究、自主发现的机会。(2)条件已知与未知的可区分性,无字证明应区分证明的给定条件与其他条件。已有研究表明,如果一个证明的待证命题与该命题成立的条件未被明确表述,那么学生将不习惯该证明。该研究提出可以利用视觉上作图的不同方式来克服此困难,并通过实验4得到证实:当区分无字证明的条件时,学生关于构造之白的补白率得到提高。(3)构造的可视性,无字证明应包含构造步骤。已有研究提出,若不考虑图形是如何产生的,则无法将结论推广至一般,且该研究也证实了当呈现更多的构造信息时,学生更容易产生严格的证明。(4)图形性质的内隐性,无字证明不应该呈现或标记学生能够推出的图形性质。教师应该培养学生“自发”的智慧,而非“接受”的智慧,该研究证实了当无字证明呈现的图形性质减少,学生更能发现并填补图形性质之白。

(5)包含人为痕迹(human agency),无字证明应该以人的数学活动形式展开。已有研究表明,抽象的、缺少人为痕迹的呈现方式会降低学生学习数学的兴趣,而包含人为痕迹的图更容易激发学生的数学参与。

同时,该研究的价值体现于三个层面。在概念层面,该研究将基于无字证明的活动定义为一种新的证明活动,以促进学生的数学学习。在理论层面,该研究开发了补白的理论框架,以分析数学证明材料以及学生对其的表现。在教学设计层面,该研究基于补白理论阐述了五项无字证明的设计原则,以促进更多学生产生更详实的证明。

最后,文章提出关于无字证明与补白的未来研究方向。一是研究与补白相关的心理过程,二是探究可以确保学生识别空白的先决条件,三是研究补白行为在合作学习过程中的传播方式,四是探究后进生如何从无字证明活动中受益。

(本文责任编辑:曾思莹;本文责任校对:曾思莹)

编者注:本刊希望作者能以中文语言习惯推介英文论文,但实际的投稿中,仍然存在大量的按英文语言习惯转译成的怪异中文,无法让读者懂得中文所表达的意思,甚至作者自己也同样不能理解。推介投稿中还常常存在关键词前后不一的情况,表明推介作者可能并不明确原文作者研究的概念,也反映了并不能充分理解原文,只是机械翻译了文章中的句子。本篇“《何时能够以少见多?探究无字证明中的补白》推介”向《学科教育研究》的作者和读者提供了一个优秀的推介撰写示范,作者刘倩雯和汪晓勤能够在完全整理英文原文的论点、论据的基础上,将作者的研究过程和研究结论用自己的语言、用符合中文语言习惯的方式表达出来,并且完全稳定了关键词的表达,能够让读者通过中文推介充分而透彻地了解英文作者的研究过程和研究结论。同时,在投稿过程中,作者对照《推介稿投稿格式规范和自查表》进行自查,保证了推介的规范性。特此向读者推荐本文,同时也推荐给本刊投稿作者参考。感谢大家对《学科教育研究》的支持,期待您提供更优质的推介稿件。



《数学评价：关于疫情时期教师实践的研究》推介

石 城, 汪晓勤

(华东师范大学 教师教育学院)



Assessment in mathematics: a study on teachers' practices in times of pandemic

Annalisa Cusi¹ · Florian Schacht² · Gilles Aldon³ · Osama Swidan⁴

Accepted: 20 June 2022
© The Author(s) 2022

Abstract

Lockdowns imposed by many countries on their populations at the beginning of the COVID-19 crisis forced teachers to adapt quickly and without adequate preparation to distance teaching. In this paper, we focus on one of the most formidable challenges that teachers faced during the lockdowns and even in the post-lockdown emergency period, namely, developing assessment that maintains the pedagogical continuity that educational institutions typically require. Based on the results of a previous study, focused on the analysis of answers to an open-ended questionnaire administered to a population of 700 teachers from France, Germany, Israel and Italy, a semi-structured interview series was designed and implemented by the authors of this paper with a small group of teachers. The transcripts of these interviews were analysed according to the interpretative phenomenological analysis methodology, with the aim of investigating teachers' own perspectives on the following: (a) the difficulties with which they had to contend, with respect to the question of assessment; (b) the techniques adopted to deal with these difficulties; and (c) the ways in which the lockdown experience could affect the future evolution of teachers' assessment practices. This analysis supported us in formulating hypotheses concerning the possible long-term effects of lockdown on modes of assessment in mathematics.

Keywords COVID-19 pandemic · Distance teaching · Formative assessment · Summative assessment · Meta-didactical transposition · Praxeologies

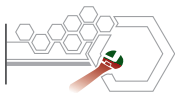
原文: Cusi A, Schacht F, et al. Assessment in mathematics: A study on teachers' practices in times of pandemic[J]. ZDM Mathematics Education, 2023, 55(1): 221–233.

原文网址: <https://doi.org/10.1007/s11858-022-01395-x>

核心概念: 远程教学 (distance teaching); 形成性评价 (formative assessment); 终结性评价 (summative assessment); 元教学转换 (meta-didactical transposition); 解释性现象学分析 (interpretative phenomenological analysis methodology)

1 引言

新冠肺炎危机开始时, 许多国家对其人口实施了封控, 迫使教师迅速适应、没有充分准备就进行远程教学 (distance teaching)。该研究关注教师在封控期间甚至在封控后的紧急时期所面临的最可怕的挑战之一, 即如何开发评价模式, 以保持学校教育通常需要的教学连续性。该文的第一作者 Annalisa Cusi 及其所在团队之前已经以开放式问卷的形式研究了 700 名来自法国、德国、以色列和意大利的教师, 进一步对其中的部分教师进行了半结构化访谈。这些访谈的记录根据解释性现象学分析方法 (the interpretative phenomenological analysis methodology, IPA) 进行分析, 目的是调查教师对疫情期间实



施教学评价的看法。该研究支持了关于封控对数学评价模式可能存在的长期影响的假设。

2 研究背景

由于学校结构的巨大变化以及新冠肺炎突发事件导致学校教学向远程教学的广泛转变，系统性的社会和教育问题变得更加明显，促使研究人员“思考新常态的潜力”。这些引发了关于这场疫情对数学教育的主要影响的讨论，比如倒退回教育更注重知识传播的风险和创造新的交流障碍。数字资源在为教师和学生创造有效的互动环境方面的潜力与其在疫情期间可能扩大存在的社会差距的风险之间的对比。通过反思新冠肺炎危机在推动数学教育的数字技术时代的作用和数学教育如何为公民提供必要的工具来面对像新冠疫情这样的全球危机，人们逐渐意识到这是一个改变的机会。

3 研究设计

3.1 理论框架

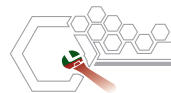
为了阐明疫情如何影响教师的评价实践，更好地描述教师在此期间应对挑战的方式，必须参考他们如何解释身处的复杂现象。该研究的理论框架为元教学转移模型（Meta-Didactical Transposition model, MDT）和内化理论，用来支持教师在专业背景下实践的描述，分析他们的实践和选择。MDT是基于人类学理论的教学法，它将数学教学解释为一种在机构内进行的普通的人类活动。根据这一理论，数学教学的特点表现在实践层面（知道如何）和标志层面（知道为何）。所面对的任务和面对任务所采用的技术是实践部分，为验证技术合理性而使用的术语及其支持元素构成了标志层面。在不同的学校环境下，由于教师生活的不同经历和相应的反思，其教学行为不断发生演变。因此，MDT认可的另一个关键理论视角是内化，以表明这一进化基础上的现象。通过内化过程，新的部分被融入现有的（教学或元教学）实践中。这一过程是由教师的经验（参与专业发展项目或合作研究项目；职业生涯的关键时期面临特殊挑战，如新冠肺炎危机）和他们从这些经验中发展出来的反思所触发的。

3.2 研究问题

通过对教师在封控时和封控后紧急时期评价实践的演变的详细分析，回答以下问题：

- 1.在封控时和封控后的紧急时期，教师在评价过程中必须应对哪些挑战？
- 2.他们是如何应对这些挑战的？
- 3.远程教学的经验在教师实施评价的实践演变方面有什么影响？

前两个问题旨在描述教师的实践特征，通过强调他们在进行评价实践的任务时必须面临的挑战（研究问题1）和教师面对这些挑战所采用的技术（研究问题2）。第三个问题旨



在通过强调内化发生（或没有发生）的不同方式以及可能的潜在原因，来描述这些实践的演变。

这三个问题具有重要的理论和教学意义：从理论的角度来看，该研究旨在阐明影响教师在评价实践中做出选择的因素，以及教师如何解释和证明他们的选择。从教育学的角度来看，该研究支持制定关于疫情封控对数学评价模式可能产生的长期影响的假设，提供可供教育家和决策者参考的设计教师专业发展方案的想法。

3.3 研究方法和研究过程

原文第三作者Aldon及其团队在2021年进行了一项国际研究。通过对700名来自四个不同国家（法国、德国、以色列和意大利）数学教师的开放式问卷回答的定性分析，发现评价是封控期间教师面临的巨大挑战，封控迫使大多数教师采用形成性评价的视角。随后的2022年，作者设计并实施了一系列对数学教师的半结构化访谈。研究对象是上述700名数学教师中自愿报名的不同学段（小、初、高）的48名教师，探讨了在封控期间和封控后，评价在数学课堂中的作用。

在采访前，参与者同意研究人员要问的问题和顺序，然后问题被翻译成受访者的语言。每次采访都从封控期间的教学和学习的介绍开始，目的是强调在封控期间的教学经验的主要变化。访谈的第二部分，重点讨论了在封控期间和封控后的评价方面。教师们被要求详细描述在进行评价时他们所面临的挑战，并反思他们的想法和评价策略的变化。表1显示了属于这一阶段的三组中心问题和目标。

表1. 教师评价实践的三组核心问题

Table 1. The three groups of central questions focused on teachers' assessment practices.

问题	问题的目标和来源
你能描述一下你在封控期间如何开展评价工作的吗？你评价学生的方式在封控后有改变吗？为什么？	这些问题指向教师在执行评价任务期间所采用的技术和证明这些技术有效性的话语。
在这段紧急时期，你对数学教学评价的观点发生了什么样的改变？	这个问题指向教师实践的标志层面元素：他们眼里的评价是什么以及面对疫情危机的挑战时评价方式发生了什么改变？
在封控期间的远程教学的经历在评价实践方面教会你什么？为什么？	这些问题指向教师对未来执行评价实践任务的实践的发展（在理论层面和标志层面）。

该研究将IPA作为数据分析的有效工具。IPA的目标是详细分析参与特殊现象的人如何感知和理解它们。主要是理解意义的复杂性，而不是测量它们的频率。因此，IPA研究收集数据的一种有效方法是进行半结构化访谈，因为它允许研究人员与研究的参与者进行对话，根据参与者的回答修改问题，探索产生的有趣想法。IPA的研究人员在数据分析过程中采用了一种迭代的方法，在可用的数据中反复分析，以了解它们的内在含义，并掌握



不同的视角。此外，IPA遵循一种特殊的分析方法，从特定的例子开始，慢慢地推导出更一般的类别。分为以下四个阶段：

第一阶段：在第一个案例中寻找主题。反复阅读一段文字记录（第一个案例），目的是注释有趣的想法，并识别与这些想法相对应的新兴主题。

第二阶段：联系主题。把第一阶段产生的主题按时间顺序列出，并以更分析化和理论化的方式重新排序，目的是理解新兴主题之间的联系。这种方法能够将一些主题聚集在一起，并识别作为高级概念出现的主题。

第三阶段：继续分析其他案例。目的是识别重复的模式，同时承认可能出现的新问题。画出上级主题的总表。

第四阶段：总结写作。这一阶段的目的是将主题转化为叙事性的讨论，穿插从文本上获得的逐字摘录来表述这个案例。

在该研究中，首先四名研究人员通过阅读访谈记录，各自进行IPA方法的第一阶段，注释与评价相关的有趣想法和涌现的新主题。此外，该研究中有趣的摘录被识别出来，并被翻译成英语，供研究小组成员分享。该小组建立了一个共享的平台，以收集和聚类不同的新兴主题，并举行了各种共同的（虚拟的）会议，以分享主题之间的初步联系。研究通过反复的会议，删减、合并，最终确定了三个主要的主题组块，每一个都对应于一个研究问题（如表2所示）。

表2. 数据分析中使用的表（节选）

Table 2. Tables used in data analysis(excerpt).

组块1		
1—A：有关学生形成性评价过程任务的挑战		摘录例证
非言语的交流被阻碍导致的困难（甚至在过渡时期）	意大利（1 LS） 法国（1 US，1 P） 以色列（1 US）	“因为学生掩盖这里（鼻子），他们还有一顶帽子，我们唯一能看到的是眼睛，感觉没有任何进展，没有直接透明的反馈，不知道你在和什么沟通。”（意大利教师，LS）
组块2		
2—A：教师应对执行形成性评价任务过程挑战的方式		摘录例证
重点关注具体活动和要求，关注问题解决	意大利（1 LS，1 US） 以色列（2 US） 法国（1 P）	“我必须改变我的教学和评价方式。我在互联网上发现了两个很棒的应用程序，它们帮助我给学生提供几何学方面的问题，并追踪他们的解决方案。该软件为解题提供了线索，学生们利用这些线索解决了问题。这个系统给我在线反馈。这让我能够在几何学上提出更多的问题，并符合学生们的理解。”（以色列教师，US）

注：LS表示初中（6-8年级或者6-9年级）；US表示高中（9-12年级或者9-13年级）；P表示小学（1-4年级或者1-5年级）



两篇摘录例证都与潜在的理论框架和研究问题高度相关。第一个摘录节选讲述了一位意大利教师所面临的挑战，他关注的是实行一个基本的形成性评价策略的困难：组织有效的课堂讨论，关注学生的理解程度。在第二个摘录，以色列教师提出她试图面对这种挑战的方式，通过描述潜在的特殊数字工具，促进形成性评价实施的关键过程，如监测学生的学习过程和给予特殊性的反馈支持他们的学习。

在第一次收集数据和确定三个主要组块之后，四个研究人员分别进行IPA方法的第二和第三阶段。在系统地分析了数据后，通过确定三个主要组块的子组块，重建新出现的主题和收集的摘录之间的联系。并对该分析过程和聚类过程的结果进行了比较。

在下面的章节中，将介绍从对教师访谈记录的分析中产生的主题，分为三个主要部分，每个部分对应于三个研究问题中的一个和已经确定的一个主要组块。即使这四个国家的教育制度所具有的一般价值观非常接近，但认定评价技术合理性的学校政策框架依然是互不相同的。因此，教师的教学行为、技术必须在学校背景下加以解释和分析。这就解释了为什么在某些情况下，几乎“相反”的想法出现在同一个子组块中。

组块1：教师在执行评价任务过程中所面临的挑战。研究者对教师访谈的分析使得能够确定以下三个主要的子组块（sub-cluster）：

- （1A）让学生参与形成性评价过程的挑战；
- （1B）让学生参与终结性评价过程的挑战；
- （1C）与实现终结性和形成性评价过程相关的更普遍的挑战。

1A-1：不同的技术问题使他们无法迅速实施典型的形成性评价策略，如设计和进行课堂讨论。教师归因于以下方面：教师和学生最初缺乏使用特殊数字工具（特别是通讯工具）方面的经验，以及使用这些新工具开展已知评价技术的方法；缺乏明确的指导方针；学生缺乏参与在线互动课程的最基本的工具如摄像头、麦克风（特别是那些经济条件不佳的学生）。

1A-2：教师追踪学生学习过程所面临的特殊困难：缺乏多模式的沟通，这使教师无法调查“学生在学习过程中的位置”，如无法通过观察学生的注视或他们的手势。

1A-3：在监测学困生的学习过程和单独辅导他们学习方面所遇到的困难：一些教师用术语“互动障碍”来描述他们在远程教学的经历。

1B-1：如何在终结性评价中防止作弊的困难。即如何确保考试的可靠性和确保学生在没有外部帮助的情况下独立完成考试。

1B-2：确定正确的评价对象方面的困难：在学生需要参加全国性考试的国家，教师



们表达了他们对缺乏明确的数学主题和评价考试的方法的担忧。

1C: 更普遍的挑战。教师与评价过程中其他重要参与者(除学生之外)的互动: 与共同设计评价任务或策略但意见不合的同事合作的困难; 与父母协同的困难, 避免测试受到干扰的困难。

组块2: 教师在执行评价任务过程中应对挑战的方式。该研究确定以下两个主要的子组块:

(2A) 如何应对学生参与形成性评价过程的挑战;

(2B) 如何应对学生参与终结性评价过程的挑战。

2A-1: 与学生及其家庭建立一种信任关系, 培养他们切实参与形成性评价过程, 促使他们开展自主学习。例如, 系统全面地与学生对话, 巩固师生关系。

2A-2: 关注具体活动和问题解决, 将其作为设计学习任务的一种方式。“让学生建造东西”, 发起解决问题的活动, “用双手和小图表解释”, 或使用应用程序来帮助学生探索特殊问题(几何学)。这些描述都暗示了使用多种表征(标志性的、象征性的)的材料来让学生参与形成性评价过程。这一策略通常与合作学习相结合。

2A-3: 追踪学生的学习过程, 调查学生的学习进度的方式。收集和检查学生的书面作业, 对教师来说是很耗时的。因而注重与学生群体进行口头交流或发起集体讨论; 关注以小组为中心的在线诊断, 例如, 使用民意调查来快速了解学生的理解和参与; 使用普通的终结性评价而不给出分数。

2A-4: 反馈。数据显示, 书面和口头反馈既使学生能够反映学习情况, 又培养他们的学习动机。在书面反馈方面, 尽管一些教师强调了“写作和评价学生的困难”, 但也需重视学生, 给予鼓励, 培养动机: 在口头反馈方面, 进行一对一(或小组)同步交流和保证学生的咨询时间。

2B-1: 应对远程教学中评价结果是否可靠的挑战。设计特别任务以了解学生是否使用特殊软件进行远程笔试, 如搜题软件解错的题目。结合了笔试和口试以检查学生是否独立完成考试。设计检查学生的真实理解的任务, 要求学生分享整个推理和论证过程(不仅仅是最终结果), 强调开放式作业在远程学习中的价值。

2B-2: 转变评价的对象。应该评价什么? 在远程教学过程中, 与其评价学生的表现, 不如只评价真正可观察到的东西(例如, 参与度、投入、坚持性、严谨性……)。

组块3: 远程教学经验对教师实践过程演变的影响。远程教学促进教师加速完成在新冠肺炎危机之前就开始的内化进程。有时由于限制阻止内化, 例如, 一些教师认为封控期



是一个“空白期”，其特征是经历与普通课堂的现实相差太远，无法内化。研究对这些反思进行了以下分类，包含6个整合后的子类，每个对应一个类别，（P）、（L）与（a）、（b）和（c）的结合：

（1）已内化的成分类型：属于教师实践层面（P）的成分；属于理论层面（L）的成分；

（2）这些成分的内化水平：（a）教师实践中现有成分的整合；（b）新成分的内化；（c）缺乏内化。

P-a和P-b子类中，教师主要参考与形成性评价概念相关的技术（新的或预先存在的）。这些技术通常涉及使用数字平台或特定的数字工具来实现不同的过程，例如：

1.使用谷歌教室收集学生的书面报告，追踪学习过程，从学生的答案出发组织课堂讨论；

2.鼓励学生与教师和同学分享材料，并对学生所分享的材料进行纠正，提供持续的反馈；

3.让学生互对答案。实现同伴评价过程；

4.通过组织学生之间的虚拟会议（如谷歌Meet或Zoom），促进合作学习；

5.设计数字测试，为学生提供即时反馈（如通过谷歌Form），以支持他们自我评价。

P-c子类中，一些教师（尤其是初中和高中教师）特意停止使用远程教学期间采用的一些技术，比如在课程外的数字平台上会见小组学生，或者在共享的文件夹中收集学生的材料并进行纠正。他们还认为专业技术缺乏内化辩护，认为专业技术对教师的要求太高。

该研究中，没有教师属于L-c，相反，大多数教师的反思都属于L-a，因为教师所描述的过程通常主要是巩固整合先前存在的关于评价的想法，这些想法已经是他们实践的标志层面的内容了。

部分教师巩固了终结性评价概念。这些教师（主要是高中教师）在远程教学期结束后会回到了他们以前“更客观”的评价方式。他们也有学校的限制，比如需要在学期结束前收集一定数量的分数，或者为学生的春季考试做准备。还有教师巩固形成性评价的视角，他们验证了应用在封控期前的许多技术。在评价过程中，通过同伴互评和自我评价的实践发现的重要性涉及所有参与者（教师、学生、学生同龄人）。评价不应该用“衡量学生的表现”或“分数”来区别学生，相反，它应该以“叙述学生的故事”为目的，一方面，支持教师适应他们的教学，另一方面，使学生意识到他们的学习。

在L-b子群中，一些有形成性评价愿景的教师描述了一个关于评价的专业思想的内化



过程:

(1) 在评价过程中关注学生情感,将形成性评价实践作为有效的工具,以防止终结性评价通常引发的消极情绪。

(2) 积极让家长参与形成性评价过程(特别是小学),以便更好地协调学生在校内外的经历。

(3) 需要协调不同层次(课堂、学校、国家)的评价实践,并将评价更多地放在数学课程上而不是教科书上。

4 结论与反思

该研究说明新冠封控对教师开展教学评价工作的影响,这改变许多教师的想法和实践。新冠危机使一些教师发现开发评价模式的其他可能方法。与此同时,这一经历也让许多教师认识到形成性评价的价值,发展或巩固形成性评价模式的重要性。数学教育需要“评价真正应该被重视的东西,而不是仅仅评价相对容易评价的东西”。

作者也提出了反思,第一,评价实践的转变是否稳定:是永久性的还是只是暂时性的?如果教师所描述的实践过程是真实内化的,那么他们在未来的实践时,也会用到这些新的技术和相应的话语。如果这种内化不是真实的,教师可能很快就会回到他们以前的评价方法。

第二个反思是制度、文化、语言环境所发挥的关键作用。教师在不同水平的学校或在不同的学校环境中也存在差异。这一点导致教师描述的实践现象存在不同的学校特征或环境特征。这些在局部上发展起来的现象,会产生更多的全球性现象吗?

教师所描述的内化过程可以比喻为在地方层面上触发的浪潮。只有政治和教育部门促进和支持稳定、更全球性的变革,这些浪潮才能得到传播。通过关注旨在深化教师专业发展的教育项目,才能支持教师实施真正的形成性评价实践。

(本文责任编辑:曾思荧;本文责任校对:曾思荧)



《美国国家地理教育委员会研究基金：因为依然会有数据的》推介

王子悦¹，徐 艳²，蒋金秀³，卢晓旭²

(1.华东师范大学 教育管理系；2.华东师范大学 教师教育学院；3.重庆师范大学 地理与旅游学院)



原文：Phil Gersmehl. The National Council's Geography Education Research Fund: Because it would still be nice to have data[J]. Journal of Geography, 2022, 121(4): 141-144.

原文网址：<https://doi.org/10.1080/00221341.2022.2127582>

核心概念：地理教育（geography education）；数据支撑（data support）；可检验的假设（testable hypotheses）

《美国国家地理教育委员会研究基金：因为依然会有数据的》这篇文章的作者是来自美国明尼苏达大学（University of Minnesota）的名誉退休教授Phil Gersmehl。原文回顾了地理教育研究领域一些代表性人物和代表性文章，呈现了教育学研究和心理学研究的差别，举出了一些老师在教学实践中做的不加控制变量的小实验的例子。作者重在指出地理教育研究领域的现实需要：缺乏数据、缺乏技术的研究性应用、解决教育问题总是从“大问题”（即宏观层面）开始，而不是基于一个个“小问题”的解决等。作者在原文中还列举出了一些可检验的假设作为研究课题以供参考。同时介绍了美国国家地理教育委员会（The National Council's Geography Education, NCGE）一个新的研究基金项目，其目标是邀请相关人员参与“评估地理教育材料和教学策略有效性”项目的研究。

作者首先回顾了1994年刊登在《地理杂志》（Journal of Geography）上的一篇文章，题为《地理教育研究的需要：如果能有一些数据就好了》，这篇文章的作者Roger Downs在摘要中表达了他的主要观点：“令人遗憾的是，地理教育领域缺乏经验数据，这些数据可能会为有关标准制定、课程设计、材料开发、教学策略和评估程序的决策提供信息和支持。”28年后，该杂志可以只做一些小改动就重印并再次刊登这篇文章，因为在今天，地理教育研究的数据需求仍然很明显。



一、心理学研究和教育学研究

Downs是宾夕法尼亚州立大学（Pennsylvania State University）地理系的一名教授，他的研究重点跨越了地理学和心理学之间的学科界限，在地理教育领域中同样具有话语权。他深刻意识到一个重要事实带来的挑战，即教育学与神经科学、心理学或视觉科学的研究有着根本的不同。

当研究对象为人类时，所有学科都面临着一个主要的限制，即研究人员必须遵循希波克拉底誓言（Hippocratic Oath）：“最重要的是：不造成伤害。”脑部扫描表明，当孩子试图解码打印出来的单词时，颞叶的某一部分可能有更多的血流量或脑电波活动，但研究人员不能在孩子的头上钻个洞来验证这一推断！同样，教育研究人员也不能强迫孩子做一些会造成自身伤害的事情。不幸的是，教育中的“伤害”可以有多种定义，比如标准化考试中的低分、州立政客或当地家长团体对孩子施加的超载压力等，都可以被认为是一种“伤害”。除了“不伤害”这一点的相似之处外，教育学和心理学的研究者有着巨大的方法论差异。一篇题为《跨越困境的桥梁：教育与认知神经科学》（Bridges Over Troubled Waters: Education and Cognitive Neuroscience）的评论文章总结了二者之间的差异（表1）。

表1. 教育学研究和心理学研究的比较

Table 1. Comparing research in education to research in psychology.

指标 Criterion	教育学研究 Research in Education	心理学研究 Research in Psychology
目标 Goal	评估/提高教学材料和方法 evaluating/improving instructional materials and methods	揭示大脑、思想、行为之间的关系 uncovering relationships between brain, mind, and behavior
方法 Method	随机选择部分对照组试验 partially controlled trials with randomly chosen groups	仔细选择匹配个体，完全控制的测试 fully controlled tests with carefully matched individuals
结果呈现 Observed outcomes	标准化考试成绩、调查问卷 scores on standardized tests, questionnaire tabulations	监控测试中的表现、大脑扫描图 performance on monitored tests, brain-scanning images
研究场所 Research setting	通常是教室 classrooms, usually as they are	经常是为适合研究而建造的实验室 laboratory, often built to suit
外部变量 Extraneous variables	无数——包括学校、教师、班级规模、课程、社会经济地位等 numerous – school, teacher, class size, curriculum, SES, etc.	尽可能减少人工干预，严密控制、严格比较 as few as possible – artificial setting, tight control, strict comparison
统计学显著性样本数量 N for statistical significance	大样本——通常是在多间大教室的100多人 usually quite large (100 s) in multiple classrooms	小样本——通常是在实验室的5-20人 usually small (5-20) in laboratory setting
可复制性 Replicability	因为很难复制相关变量，可复制性很低 low, very difficult to duplicate all of the relevant variables	如果调查人员正确记录程序的话，可复制性很高 theoretically high, if investigators document procedures properly



新技术能改变研究：当扫描技术被最终广泛使用时，改变了人们研究人类大脑活动的方式。更快的计算机和互联网连接虽然使考试的管理、评分和制表变得更加容易，然而，与神经科学和心理学中大脑扫描技术的影响相比就相形见绌了。

二、“大观念”的吸引力

地理教育工作者应该意识到所有研究性学科感染的一种“病毒”：声誉（通常也是物质上的）和奖励往往会流向那些解决宏观层面“大问题”，而不是基于细节研究“小问题”的人。这种对“大观念”（big ideas）的偏好造成了一种广泛流行的现象，即解决教育问题的方法总是“从顶层开始”，即编写新标准、重组课程、重新设计整个班级或做出其他重大变革。重大变革的提议通常都有一个有效的真理作为核心，但人们太容易忽视 Arthur C. Clarke 的一句话，“在原始数据意义上的信息不是知识，知识不是智慧，智慧不是远见，原始数据意义上的信息是实现这一切的第一个必要步骤。”

这种对“大观念”的偏好在没有严重实际后果的情况下，仅仅是一种学术上的抱怨而已，但它往往会影响那些提供研究资金的人。因此，地理教育研究项目的资助者应该认识到这种偏好的问题，并确保为那些寻求原始数据意义上的信息的研究者提供足够的支持，因为这有助于回答重大问题。

三、举例：地理和阅读之间可研究的问题

作者又回顾了2004年，这一年是地理教育研究史上的又一个里程碑。此时，大脑扫描技术已经成熟，且价格便宜。Committeri 等将一份重要的研究报告发表在高度专业化的《认知神经科学杂志》上，报告的标题是《空间认知的参考框架：不同的大脑区域以观察者、物体和地标为中心参与判断目标位置》（Reference frames for spatial cognition: different brain areas are involved in viewer-, object-, and landmark-centered judgements about object location）。紧随其后，有大量类似的研究文章发表，这些研究包括人类大脑是如何接受、处理和交流空间信息的。2009年，Stanislas Dehaene 在《大脑中的阅读》（Reading in the Brain）一书中发表了一篇研究总结，其中给出了一段令地理学家感兴趣的总结，“视觉空间注意对阅读的正常发展至关重要。良好的解码技能不仅来自字母的形状和语音之间的联系，字母还必须以适当的方向，适当的空间位置和正确的左右顺序被感知。在年轻读者的大脑中，必须在识别字母和单词的腹侧视觉通路和背侧视觉通路之间进行协作，后者会为它们在空间中的位置编码，并规划眼球运动和注意力来辅助阅读。当这些角色中的任何一个出错时，阅读就会一败涂地。”

在这一系列空间概念中：形状、方向、相对位置、左右顺序、眼球运动和空间注意力，我们可以选择一个作为地理教育研究小课题的例子，研究它是否可能会对小学阅读教



学产生影响。例如，Dehaene用了36页，超过《大脑中的阅读》10%的篇幅，来研究方向感的视觉感知及其与正常阅读和某些类型的阅读障碍可能存在的联系：在生命的早期，几乎所有的孩子都会在阅读和写作中犯镜像错误，跨对称视图进行泛化的能力也确实是视觉系统的基本能力之一，当孩子学习阅读时，他们必须“忘记”镜像泛化，以便将“b”和“d”视为不同的字母。

作者提到，这些年来，有几位老师一直在和自己的学生一起做绘制教室地图的活动，其中还使用Simon says游戏（一款英国传统的儿童游戏）和一些脚手架活动。根据这些老师的记忆发现，学生在写作中颠倒字母的次数减少了。

作者对此做出解释：这是合理的，因为这与现今理解的大脑发育是一致的。不幸的是，像这样的观察数据都是奇闻轶事，如俚语所说，“很多奇闻轶事聚在一起就成了数据”；用科学术语来说，这个例子中有太多不受控制的变量：每年的班级都有不同的“个性”，可能会以不同的方式或不同的顺序来教授主题，包括返校节橄榄球赛、大风暴或流感爆发（更不用说像新冠这样的大事件）等事件都会对其产生影响。教师的记忆更是一个复杂的东西，也有许多令人混淆的影响因素，包括众所周知的锚定和确认偏差¹。简而言之，“有一些数据就好了”。

四、地理教育中可检验的假设

方向感的地理课能帮助学生克服视觉系统在阅读和写作时忽视字母方向和左右顺序的自然倾向吗？对这个问题（以及许多其他类似问题）的肯定回答将对课程讨论非常有帮助。最近的教学趋势是减少科学、社会研究或艺术等科目的时间，以便腾出时间给进行标准化考试的阅读和数学更多的指导。哈莱姆（Harlem）一所公立学校的七名老师曾经尝试过一项不加控制的“实验”，他们在幼儿园和一年级每周增加40-80分钟的地理课，结果显示再次进行标准化考试的阅读和数学的分数要高得多。

地理课对阅读成绩可能产生的有益影响还可以重新表述为一项可检验的假设：

背景：如果学生学习一些方向感更显著的课程，那么学生在写作中的字母颠倒情况可以减少。例如，通过方向命名教室墙壁，将方向名称用于教室管理任务，如排队去吃午饭，玩Simon says游戏（其中要包括指向、面对或根据命令在特定方向上移动的活动）以及制作桌面、教室、操场的模型和地图。

问题：对于像这样注重方向感的地理课来说，时间安排重要吗？具体来说，如果在幼儿园、一年级早期或一年级晚期进行，会导致字母颠倒错误期提前结束吗？

1 锚定是一种重要的心理现象，是指当人们对某个事件做定量估测时，会将某些特定数值作为起始值（参照物），起始值（参照物）像锚一样制约着估测值。同样，在做决策的时候，会不自觉地给予最初获得的信息过多的重视。



这样可检验的假设还有很多，作者在原文中举出了以下的五个例子：

1.需要多少个平面方位词的练习活动才能提供一个好的支架来帮助学生定位目标物呢？

2.在地图上需要多少背景信息来帮助二年级的学生记住学校、商场或其他熟悉地点的位置？四年级和六年级的答案会不同吗？

3.用我们的数字国家地图集进行什么样的顺序活动才能在脑海中产生最持久的国家地区形象？

4.在比较早期人类迁徙、历史帝国、世界宗教或气候变化影响等主题时，哪种地图投影最有效呢？

5.哪种旅行前（或旅行后）的写作活动能从野外实习中获得最持久的学习呢？

五、美国国家地理教育委员会研究基金

为了支持这类（以上提出的可检验的假设例子）研究，像NCGE这样的组织至少可以部分满足Downs曾关于地理教育研究而列出的四个要求：一种对实证研究持支持态度；以一些基础研究为例；一个具体的研究议程；以及一种传播成果的方式。

值得庆幸的是，NCGE启动了一个新的研究基金项目，以“评估地理教育材料和教学策略有效性”为主题，邀请相关人员参与研究。这项研究的主要重点应该放在为了解决大问题而必须回答的“小问题”上。委员会可以资助来自高校地理院系、州地理教育联盟等机构中的教师团体、教师个人或研究生个人。该研究基金的具体管理细节在作者撰文时仍在起草中，NCGE计划在2022年10月份于明尼阿波利斯召开一次特别会议。会议的目标是征求意见，并开始邀请申报，公布分配资金和评估方案。感兴趣的各方可以提出该基金项目可能支持的研究类型的建议或可检验的假设例子，这将会被公布在NCGE网站上。

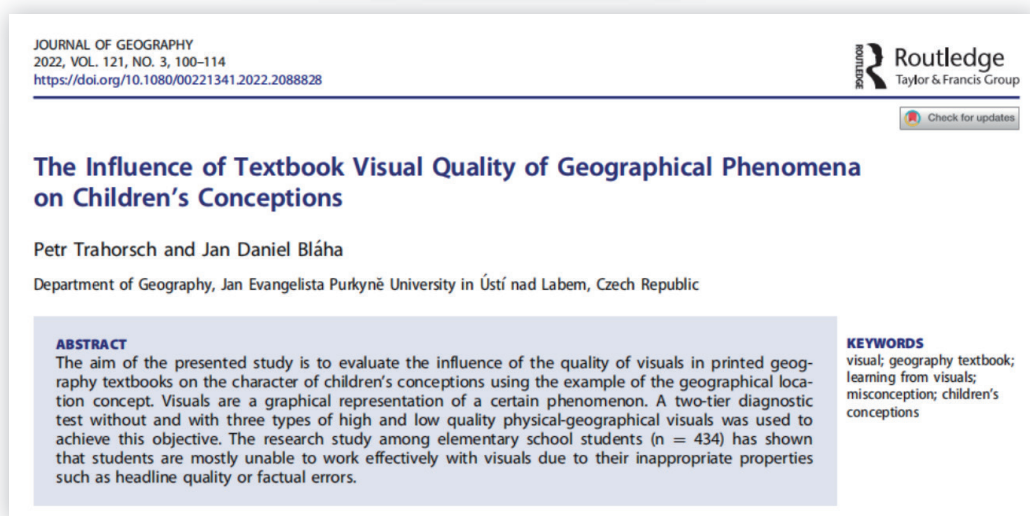
综上所述，该论文中作者首先从Downs的文章引入，到心理学研究和教育学研究的对比、再到一些老师们开展的不加控制变量的教学小实验。这强调了这样一个事实：涉及教育学的过程中有太多不可控制的因素来影响结果的判断，所以基于实证研究的数据必不可少，同时发展中的技术也应该在教育学研究上得以应用。其次，对于“大观念”研究的偏向，作者也没有否认其存在的现实意义，但是作者更倾向于地理教育研究要以解决“大问题”而不可避免的“小问题”为出发点、为重点。作者在回顾具有的里程碑意义的研究中，再结合现实教育的情况，列举了一些可检验的假设作为地理教育领域值得研究的课题以供参考。最后，作者介绍了NCGE新的研究基金项目提供了一个平台，以收集值得研究的课题建议和可检验的假设例子。

（本文责任编辑：曾思莹；本文责任校对：曾思莹）



《教科书中地理现象视觉资料质量对儿童概念的影响》推介

安爱玲, 朱瑾烨, 卢晓旭
(华东师范大学 教师教育学院)



原文: Petr Trahorsch, Jan Daniel Bláha. The influence of textbook visual quality of geographical phenomena on children's conceptions[J]. Journal of Geography, 2022, 121(3): 100-114.

原文网址: <https://doi.org/10.1080/00221341.2022.2088828>

核心概念: 地理教科书 (geography textbook); 视觉资料 (visuals); 可用性 (usability); 从视觉资料中学习 (learning from visuals); 儿童概念 (children's conceptions); 误解 (misconception)

近几十年来,随着教科书的创作、编辑到出版都与现代信息技术紧密相关,地理教科书中视觉资料 (visuals) 的数量和质量也越来越重要。地理教育中的视觉资料可以理解为某种现象或过程的图形表示,人们能够更快、更准确地将注意力导向相关数据,从而对 (视觉) 地理信息的空间方面进行更快地认知处理。对于视觉资料质量的评估可以用可用性 (usability) 这一指标,包括清晰度、可读性、专业性等。

儿童脑中的概念表达了个体对外部世界的理解和解释。视觉资料和其他媒体 (如视频、文本、3D模型、地图集等) 可以被认为是影响儿童对地理现象概念理解的外部因素。很多研究及多媒体学习理论 (multimedia learning theory) 表明,某些情况下,视觉资料可以提高学生学习效率和对地理现象的理解水平。其中,视觉资料的心理教学质量 (psychodidactic quality) 是对学生学习地理现象相关知识时的认知过程产生积极影响的



关键。但如果视觉资料的表达方式或内容不当,则会降低其质量,影响学生的学习效果,影响对地理现象的认知。其他科学学科类似研究主题(视觉化和儿童概念间的关系)的研究结果也表明,教科书中的视觉资料问题有可能导致儿童产生错误的概念。

然而,仅凭现代技术不足以创造出高质量的视觉资料,需要以科学研究为前提。因此,从地理教育的需要出发,《教科书中地理现象视觉资料质量对儿童概念的影响》一文以地理位置概念为例,旨在通过任务解决来评估地理教科书中照片、地图、示意图等视觉资料的质量对儿童概念的影响,同时确定各种类型的视觉资料对儿童概念影响的差异。根据研究目的,原文作者Petr Trahorsch等提出了主要的研究问题:以地理位置为例,不同的视觉资料质量对中小学生学习儿童概念的形成有什么影响?并基于该研究问题,提出了以下两个假设:

假设1:使用高质量视觉资料的学生比使用低质量视觉资料的学生能取得更好的成绩。即假设具有可用性特征的视觉资料可以帮助学生解决问题,从而有助于消除概念误解。

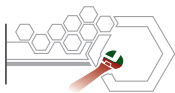
假设2:学生在有视觉资料测试时的测试成绩将高于没有视觉资料测试时的成绩。即假设视觉资料可以帮助解决问题,从而有助于学生的认知过程。

本文所推介原文的作者为Petr Trahorsch和Jan Daniel Bláha,两人都在捷克的Jan Evangelista Purkyně大学地理系任教。研究在捷克西北部的6所公立学校进行。参与者是来自两个年龄段的434名学生:9-11岁的小学生221名,13-15岁的初中生213名。特意选择这两个年龄段是因为他们是教科书中视觉资料理想的使用者,且他们教科书中所涵盖的主题非常相似。

研究工具及应用

研究选用双层诊断测验(two-tier diagnostic test)方法来测试视觉资料对儿童概念的影响。测题由答案选择层和验证层组成,第一步测试中,学生需要在3-5个备选答案进行选择,并在第二步的开放测试项目中证明他们的答案(具体测试项目的测题详见原文附录)。作者从他们曾经依据视觉资料质量(可用性)标准对捷克现行地理教科书中的视觉资料评价的结果中,选出了高质量和低质量视觉资料作为测试用的视觉资料。

对于地理位置这一概念,考虑到学生年龄组所学课程具体现状,研究将地理位置概念具体到了地貌学和水文学两个主题领域中。并根据对教科书中教学内容的分析,从每个主题领域中选择了三个与地理位置概念相关的不同层次的自然地理现象(physical and geographical phenomena):地貌学的主题以山峰(peak)、地表等级(classes of the land surface)、坡度(slope angle)为代表;水文主题以河口(river mouth)、河段等级(classes of river segment)和左右支流的区别(the distinction between right and left river



tributaries) 为代表。选择自然地理而非人文地理,是因为在捷克中小学教学中前者应用范围大于后者,且自然地理现象随时间推移具有较强稳定性,以及自然地理术语逻辑系统更为精确。

两个主题下的测试都由无视觉资料测试项目和有视觉资料测试项目(有照片、有图示、有地图)两部分组成,并且覆盖了各自主题中三个被测试的地理现象,有视觉资料测试项目中都设置有高质量和低质量的视觉资料。为确保研究工具的可靠性和有效性,研究采用了专家分析法,并在研究正式开始前进行了两次试测。根据专家意见和试测结果对测试中发现的问题进行了修改。

正式研究中测试分为三个阶段(见图1)。第一阶段(前测阶段),所有受试者都接受了无视觉资料的地貌学和水文学概念测试,以诊断受试者现有知识和对被监测的地理位置现象的潜在误解。该阶段的结果可以成为判断视觉资料在完成任务中是否发挥作用的基础。第二阶段,所有受试将被随机分为两组,一组接受高质量的地貌视觉资料教学性测试(didactic test),另一组接受低质量的地貌视觉资料教学性测试。第三阶段是以水文学为重点的测试,完成了高质量地貌视觉资料教学性测试的一组在此阶段接受低质量水文视觉资料教学性测试,另一组反之。整个测试时间不作限制,但学生通常在25分钟内完成一次测试。

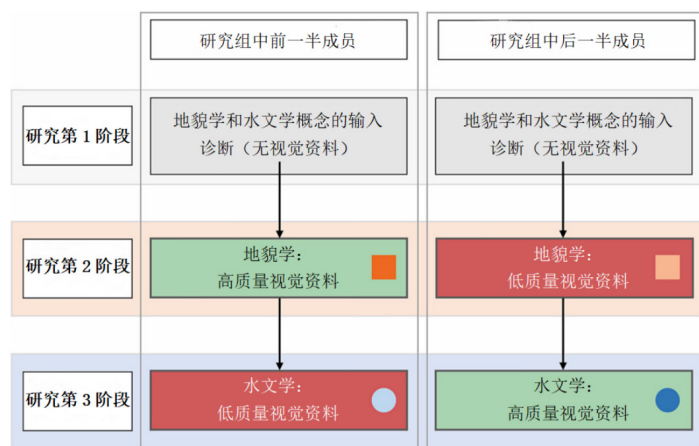
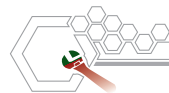


图1. 用于检验视觉资料质量对儿童概念影响的模型

Figure 1. Model used to test the influence of the quality of visuals on children's conceptions.

数据分析及研究发现

数据分析时,研究对测试中的多项选择题答案进行了频率统计,而对于开放性测题的回答,则由经过培训的2名地理研究生作为编码员进行编码分析,并在编码后对整体结果和答案进行综合评分。试测阶段显示,两位评价者的任务评分具有一致性(Cohen Kappa系数达0.76)。由于正式测试中测试结果数据未通过正态性检验,呈非正态分布,故选用



非参数统计检验方法，同时结合显著性检验和效应大小计算对定量数据进行分析。效应大小计算（effect size calculation）的结果即效应量，可用Cohen's d值表示，它可以量化变量之间的关联程度，效应量越大，两个变量之间的关系越强。与显著性检验不同，效应量不受样本限制。

研究首先用示范性例子对测试项目中具体情况进行了统计分析。

根据对前测阶段的无视觉资料测试结果具体分析发现，所有两组学生全部概念的效应量均小于0.2（见表1）。且除“左右支流的区别”概念外，其余概念测试结果P值均大于0.05，不存在显著差异。总体而言，随机分成的两组学生对地理位置概念的认知水平实质上相同。

表1. 无视觉资料的各测题测试结果

Table 1. Results of individual test items without visuals.

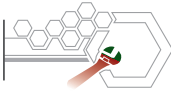
测题序号	所测概念	P值	效应量	正确率(%)
第1题	山峰	0.11	0.08	71
第2题	地表等级	0.35	0.05	52
第3题	坡度	0.15	0.07	66
第4题	河口	0.92	0.01	18
第5题	河段等级	0.32	0.05	37
第6题	左右支流的区别	0.02	0.12	42

从后续阶段的有视觉资料测试具体结果中（见表2）发现，在正确率方面，水文学主题概念的定位正确率较低，学生没有从学校的教学中充分掌握这些概念，高质量的视觉资料也不能帮助他们理解这些概念。

表2. 有高质量和低质量视觉资料的各测题测试结果

Table 2. Results of individual test items with high and low quality visual.

测题序号	视觉资料类型	所测概念	P值	效应量	高质量视觉资料组	低质量视觉资料组
					正确率(%)	正确率(%)
第1题	照片	地表等级	<0.01	0.30	37	67
第2题	照片	山峰	0.01	0.13	91	97
第3题	照片	坡度	<0.01	0.28	51	23
第4题	图示	山峰	0.34	0.05	87	83
第5题	图示	地表等级	<0.01	0.19	83	67
第6题	图示	坡度	0.01	0.14	69	55
第7题	地图	地表等级	<0.01	0.66	08	73
第8题	地图	山峰	0.87	0.01	62	61
第9题	地图	坡度	0.03	0.11	40	51
第1题	照片	河段等级	0.58	0.03	33	36
第2题	照片	左右支流	<0.01	0.18	55	37
第3题	照片	河口	<0.01	0.37	67	30
第4题	图示	河段等级	<0.01	0.25	54	29
第5题	图示	左右支流	0.28	0.05	47	37
第6题	图示	河口	<0.01	0.23	43	22
第7题	地图	河段等级	0.31	0.05	18	14
第8题	地图	左右支流	0.11	0.08	27	20
第9题	地图	河口	0.26	0.06	62	56



在效应量方面，重点关注极高值的结果，因为其对地理教学具有重要性且在实践中具有最大程度的可用性。在地貌学视觉资料测试中，第7题效应量为0.66。结合学生在开放性测题中的回答发现，该题视觉资料中（图2），由于图例和色彩对比度的原因，学生在使用低质量视觉资料答题的情况比使用高质量时好很多。这表明地图构成要素是对成功解决该问题并对地理现象进行定位的关键因素。在水文学视觉资料测试的第3题中，高质量照片（图3）中简明标题清楚地描述了照片的各个要素，这对学生解决该问题起到了帮助。

学生对某些问题的回答还表明视觉资料的质量会对误解的来源产生影响，方案的复杂程度和存在事实错误会影响学生答题的正确率（具体示例详见原文）。

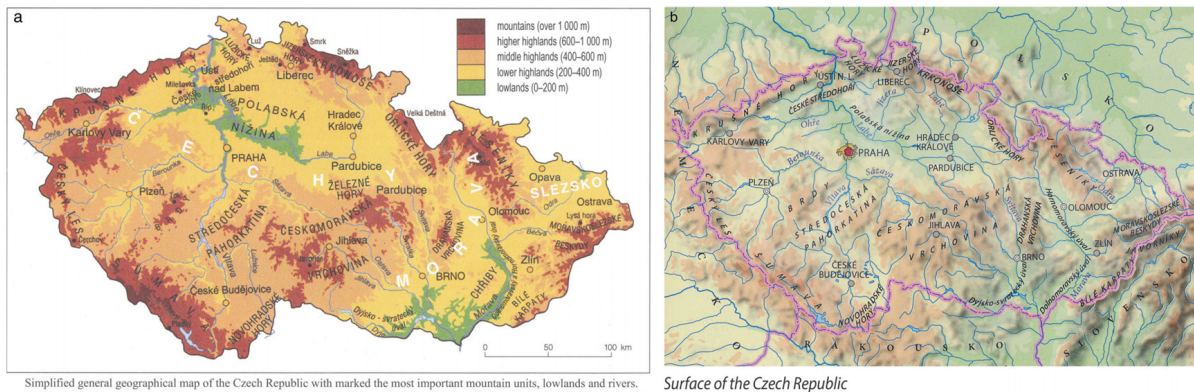


图2. 低质量(a)和高质量(b)地貌图示例

Figure 2. An example of a low-quality (a) and a high-quality (b) geomorphology map.

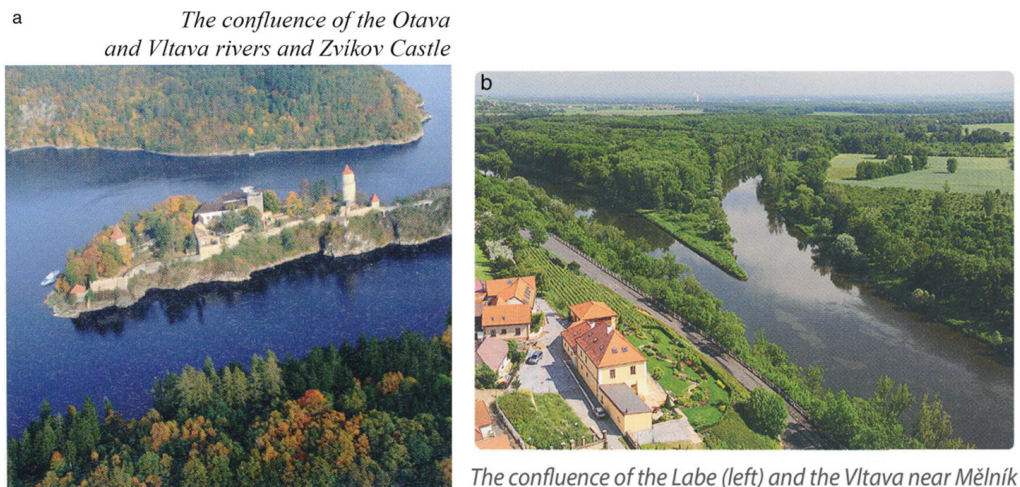


图3. 低质量(a)和高质量(b)水文照片示例

Figure 3. An example of a low-quality (a) and a high-quality (b) hydrology photograph.

简而言之，视觉资料可能会影响儿童对地理位置概念的认知，主要表现在视觉资料的个别属性上，比如视觉资料中的标题和事实性错误、关键的地图组成元素和配色都可能造

成负面影响。

最后，研究对测试数据整体进行统计分析。结果对研究最初提出的两个假设进行了回应：

(1) 由于在无视觉资料的前测阶段中两组学生正确率没有统计学上的显著差异(图4; 地貌学: $P=0.09$, $d=0.17$; 水文学: $P=0.06$, $d=0.24$), 说明两组学生对地理位置概念的初步知识掌握程度相同。有视觉资料测试阶段分析结果表明, 在水文学主题中, 使用高质量水文视觉资料的学生比使用低质量水文视觉资料的学生取得更好的效果($P=0.01, d=0.62$), 表明水文视觉资料质量对测试结果具有中等意义的影响。但在地貌学主题中, 使用高质量和低质量视觉资料的两个学生小组间不存在统计学差异(图4; $P=0.66$; $d=0.04$)。因此, 无法证实学生在使用高质量视觉资料下学习的表现更好, 即假设1无法证实。

(2) 在对个别学生无视觉资料和有高质量或低质量视觉资料测试成绩的差异进行个案分析发现, 无论使用的视觉资料质量如何, 使用了视觉资料测试后学生(除一个例外)的成绩基本都有所下降。在水文学主题下, 地理教科书中的高、低质量视觉资料都可能对学生知识系统产生非常消极的影响。仅有在地貌学主题下, 使用了高质量地貌视觉资料的结果比无视觉资料的结果有所改善。因此, 无法证实学生在使用视觉资料时的测试成绩高于没有使用视觉资料测试时的成绩, 即假设2无法证实。

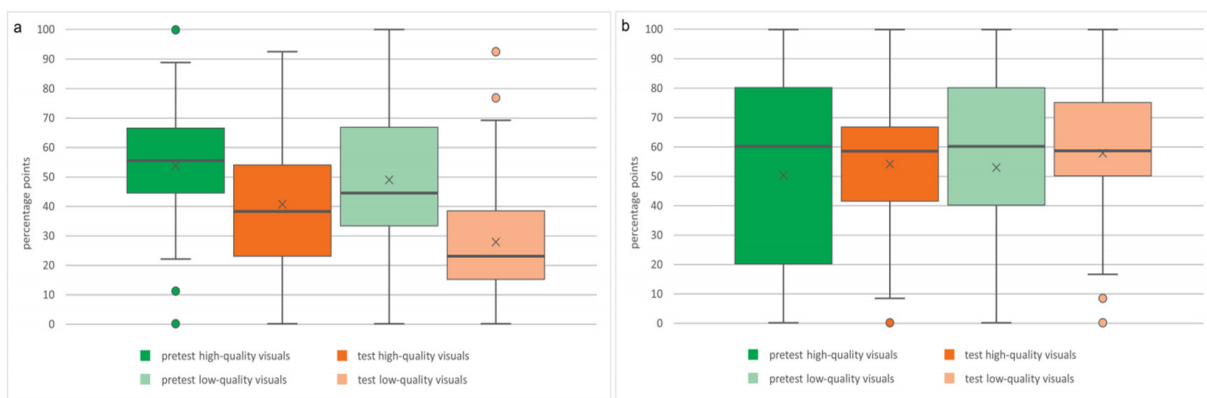


图4 使用高质量和低质量视觉资料的学生小组的正确率比较(水文学: a; 地貌学: b)

Figure 4. Comparison of the success rate in groups of students working with visuals of high and low quality (hydrology: a; geomorphology: b).

此外, 研究还发现无论视觉资料质量如何, 视觉资料的类型会对测试的正确率产生显著影响(图5; $P < 0.01$)。学生使用高质量的地貌学示意图(schema)和使用低质量的地貌学地图(map)时的正确率都相对较高。

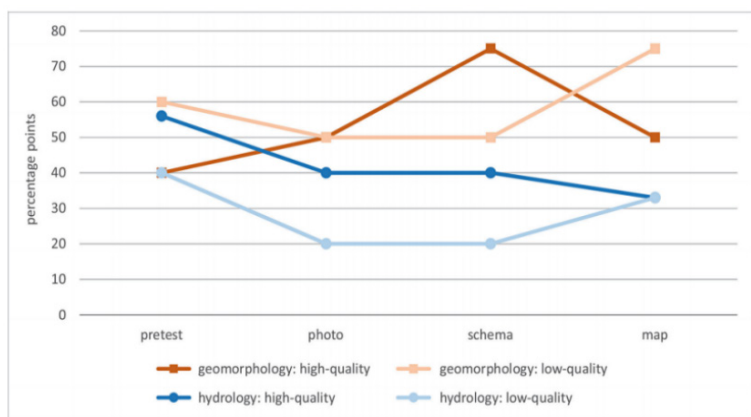


图5. 个别测试中不同视觉资料类型下正确率的中位数的变化

Figure 5. The development of a median in individual tests depending on the type of visual.

综上所述，地理教科书中使用的视觉资料有可能成为概念误解的来源，但是否出现错误取决于视觉资料的个别属性及其需解决的学习任务的概念。正是这些缺点会导致学习效率的降低和概念的误解。这项研究在一定程度上验证了不同领域中前人的相关研究成果，且表明要批判地看待视觉资料对学习过程的有显著积极影响的假设。

不过，该研究同样存在一些不足，主要可以概括为四点：（1）样本选择并非完全随机，可能在一定程度上降低了研究结果的代表性；（2）研究主要采用的定量设计可能在诊断儿童概念时的结果较为浅显，应结合定性设计的测题来补充，以解释学生使用视觉资料时所用的思维策略，理解视觉资料与学生内心的互动；（3）儿童概念的影响因素较多，从而很难确定视觉资料就是导致概念误解发生的关键因素；（4）测试任务中的措辞表述可能会对测试结果产生干扰。在今后的研究中，可以针对这些不足进行改进。

从地理教学的角度，作者认为应该提供含有帮助学生学习的视觉资料，且视觉资料应由专业的教材编写者团队（如专业地理学家、地理教学专家、平面设计师、制图师等）创建。地理教师需要把握好不同类型视觉资料的具体特征，在日常地理教学中积极帮助培养学生的识图能力，将其发展为学生有效地处理视觉资料的基本技能，以达到较好的教学效果。

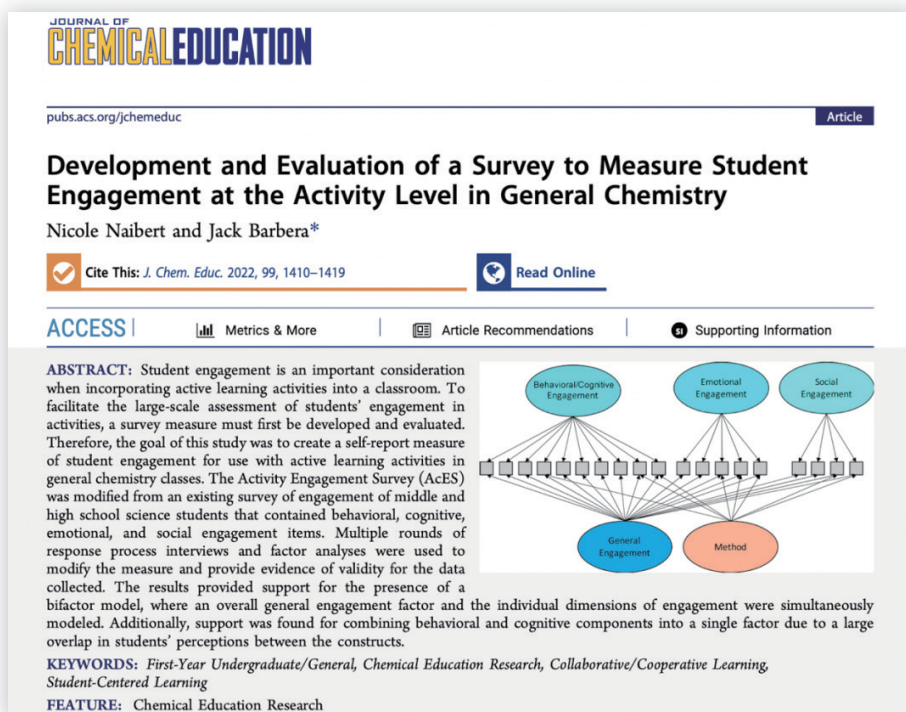
（本文责任编辑：曾思莹；本文责任校对：曾思莹）



《一项测量化学通识课程上学生活动参与度调查的开发评估》推介

皇甫倩, 邓欣雨

(西南大学 教师教育学院)



原文: Naibert N, Barbera J. Development and Evaluation of a Survey to Measure Student Engagement at the Activity Level in General Chemistry[J]. Journal of Chemical Education, 2022, 99(3): 1410–1419.

原文网址: <https://doi.org/10.1021/acs.jchemed.1c01145>

核心概念: 本科一年级 (first-year undergraduate); 化学教育研究 (chemical education research); 合作学习 (collaborative/cooperative learning); 以学生为中心的学习 (student-centered learning)

主动学习 (active learning) 的基本目标之一, 是通过为学生提供直接参与课程材料的机会, 将学习的重点从教师转移到学生身上。由于学生的参与度 (engagement) 被认为是一种可塑的状态, 很容易受到周围环境的影响, 因此, 了解学生如何在课堂上参与个人活动的机制有助于明晰不同类型的学生所适用的不同活动。已有研究证实了学生的高课堂参与度与学生的高学业成绩相关, 据此推测, 对特定活动的参与度越高可能是提高学生对相应的学习活动的理解程度的一个促进因素。此外, 学生的参与度也被纳入与动机相关的模型中。为了便于对学生参与活动的程度进行大规模的评估, 必须首先制定和评估一个调查措施。

学生的参与度通常被认为是一个由多个维度组成的元结构。这些因素通常包括行



为、认知和情感参与。行为参与 (behavioral engagement) 通常被定义为学生在课堂上的参与。对于一个特定的活动, 这可能包括诸如提问、参与等行为。认知参与 (cognitive engagement) 通常被定义为学生在学习中的投入, 包括试图理解他们的错误, 并试图将新想法与以前学到的概念联系起来。情感参与 (emotional engagement) 是指学生对周围环境的情感反应, 包括他们的同伴、老师和活动本身, 它可以由许多不同的情绪组成, 如价值、无聊和兴趣。参与度理论上属于一个元结构, 因此同时评估参与的三个维度可以更深入地了解这些维度是如何相互关联以及影响结果的。除了参与的这三个维度外, 参与的社会方面也被概念化为与STEM学生相关的参与定义的重要组成部分。社交参与 (social engagement) 指的是为学生与他人的互动以及建立关系的意愿。在考虑认知参与时, 协作性的社会互动也被作为一个定义的组成部分, 当学生被指导与他人合作时, 这可能尤其重要。

多种方法已经被用来衡量学生在课堂上的参与度。然而, 由于这些维度的内在性质, 自我报告的调查方式被提倡用来同时收集关于参与的各个维度的数据, 也允许为每个学生收集数据。至今没有一项自我调查同时评估行为、认知、情感和社会参与。

在将主动学习活动融入课堂时, 学生的参与度是一个重要的考虑因素。为了促进对学生参与活动的大规模评估, 必须首先制定和评估调查措施。因此, 该研究的目的是创建一种学生参与度的自我报告测量方法, 用于化学通识课程中的主动学习活动。活动参与度量表 (activity engagement survey, AcES) 是从现有的对初中和高中理科学生参与的调查中修改而来的, 包括行为、认知、情感和社交活动项目。多轮反应过程访谈和因素分析被用来修正测量, 并为数据收集提供有效性证据。这些结果为双因素模型的存在提供了支持, 同时在其中对总的一般参与因子和个人维度参与因子都进行了建模。另外, 支持将行为和认知成分合并为一个因素, 因为在构建过程中学生对这两者的看法有很大的重叠。

《一项测量化学通识课程上学生活动参与度调查的开发评估》研究旨在开发一种测量学生活动参与度 (AcES) 的方法, 适用于通识化学课堂上所开展的活动。这项研究的作者Nicole Naibert和Jack Barbera均来自美国波特兰州立大学的化学系。作者修改了Wang等学者最初为中学生开发的参与度量表, 并收集了原量表反应过程和内部结构效度的证据, 其中包括评估最能代表参与的多维本质的不同因素结构。该研究主要探究2个问题: (1) 修改后的学生参与度测量工具对我们的学生群体和主动学习环境的影响如何? ; (2) 哪种因子结构最能代表测量数据?

该研究的数据在该校2020-2021学年连续3个学期的通识化学 (general chemical, GC) 课程上收集的。所有的课程都是以同步远程的形式教授的, 由讲座日和活动日两个部分组成。讲座日由教师授课, 共有三名讲师授课参与 (A、B、C), 每个讲师开展活动



的方式是相似的，因此可以视为同质性。活动日以工作表活动（worksheet activities）为中心，工作表包含关键问题、练习题和某一个化学主题模型的相关问题，类似于指向过程的引导探究学习材料（process-oriented guided-inquiry learning, POGIL）。该学年共9次活动，每个学期3次，在秋季学期，学生以每组3-7名的形式被随机分配到Zoom分组讨论室来完成前两项活动。第三次活动，学生们可以自愿选择学期末分组讨论室的类型，包含（1）打开视频后，与他人合作，（2）关掉视频后，与他人合作，（3）在休息室独立工作并能够询问学习助理（learning assistant, LA）问题，或（4）在大教室完全独立工作。选择与他人合作的学生（即选项1或2）这学期被分配到同一个分组讨论室，每个房间大约有四名学生，所有学生都要完成工作表。在冬季和春季学期，6次活动均采用四种分组讨论室的模式。

表1. 每一学期化学通识课程信息

Table 1. Course information for each general chemistry (GC) section and term.

term	instructor	week-1 enrollment
fall (GC I)	A	N = 297
	B	N = 332
winter (GC II)	A	N = 190
	B	N = 266
spring (GC III)	A	N = 186
	C	N = 156

学生活动参与度的问卷调查包含了学生在工作表活动中的行为、认知、情感和社会参与相关的项目，改编自Wang等学者的研究，共26道题，其中7道行为参与、6道认知参与、7道情感参与和6道社会参与，均采用李克特六级评分法。每个学期末（即3次活动结束后）进行1次调查，最终得到1个包含3次调查、覆盖9次活动的总数据库，只有那些同意接受调查并表示他们在课堂上完成了一项活动的学生的回答才会被纳入研究数据。通过基于主题的检查项和基于响应的检查项来清理数据，共有8%的回复被删除，最终保留了1287条回复。调查结束后，邀请学生参与访谈，共21名学生参与了访谈，第一学期13名，第二学期8名。所有的访谈都是通过Zoom远程完成的，并配有音频和视频记录。在访谈过程中，首先通过Zoom中的聊天功能为学生提供了一份空白调查副本的链接。学生们被要求回想活动，然后再次填写调查问卷。在他们完成后，他们被要求大声朗读每一项，说出他们选择的答案，然后解释为什么他们选择了他们所做的反应。如有必要，研究人员会询问后续问题，以获得更多细节或澄清。两位评分员独立对访谈视频进行分析，存在争议的部分将由二者讨论商定，访谈的定性数据主要是为了消除问卷调查的定量数据中可能混淆的内容。

采用验证性因子分析（confirmatory factor analysis, CFA）来评估不同的模型，旨在



通过拟合统计数据以确定最优模型。不同模型的拟合结果如表2所示，用行为（B）、认知（C）和情绪（E）来表示不同因素。经过一些修改后，每个量表都具有良好的内部结构效度和良好的信度。最终的19项的活动参与度调查（ACES）包括5道行为参与、5道认知参与、5道情绪参与和4道社会参与。

表2. 不同模型与整个数据集（个人和社会学生）的数据模型拟合统计（N=1248）

Table 2. Data-model fit statistics for different models with the entire data set (individual and social students) (n=1248) .

model	χ^2 (df)	p-value	CFI	TLI	RMSEA [90% CI]	SRMR
unidimensional	1508.221 (90)	<0.001	0.770	0.731	0.136 [0.130-0.142]	0.095
B-C-E correlated	495.374 (87)	<0.001	0.937	0.924	0.072 [0.066-0.078]	0.051
BC-E correlated	518.093 (89)	<0.001	0.935	0.923	0.073 [0.067-0.079]	0.052
BC-E bifactor	377.243 (75)	<0.001	0.955	0.937	0.066 [0.060-0.073]	0.037
BC-E bifactor with negative method factor	244.350 (69)	<0.001	0.974	0.960	0.053 [0.046-0.060]	0.030

^aBold values indicate the results met the suggested criteria based on recommendations from Hu and Bentler.²⁶

1.单维模型：AcES中15项相关（如图1所示）。

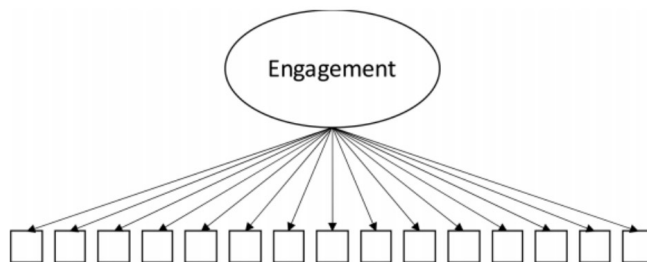


Figure 1. Unidimensional model of engagement, where all 15 AcES items are related to a general engagement factor.

图1. 一维模型

2.三因素（B-C-E）模型，其中三因素参与是独立的，但相互关联（图2）。行为和认知因素之间的因子相关性很高（0.953），高因子相关性可以表明这两个因子测量的是相同的结构。

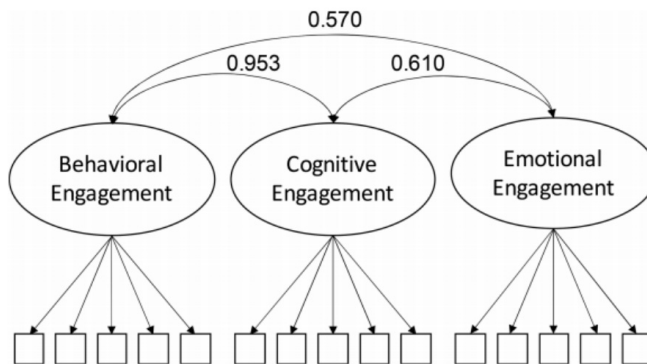


Figure 2. B-C-E correlated model where the individual behavioral, cognitive, and emotional engagement dimensions are correlated. Standardized factor correlations found through CFA are included.

图2. B-C-E模型

3. BC-E模型：将行为和认知因素合并为行为/认知（BC）因素。

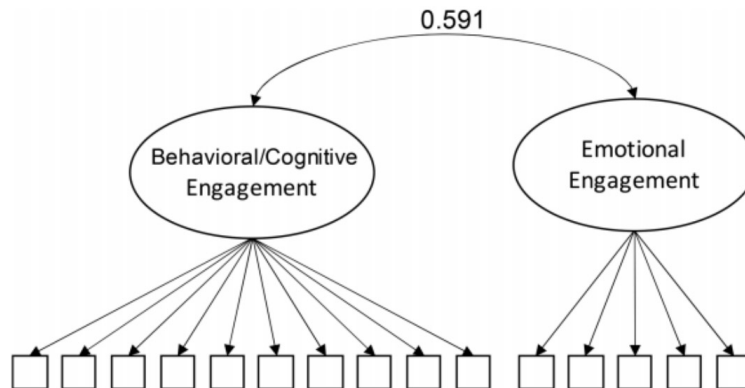


Figure 3. BC-E correlated model with a combined behavioral/cognitive engagement factor and individual emotional engagement factor. The standardized factor correlation found through CFA is included.

图3. BC-E模型

4. 双因素模型：BC-E模型（15项AcES）。

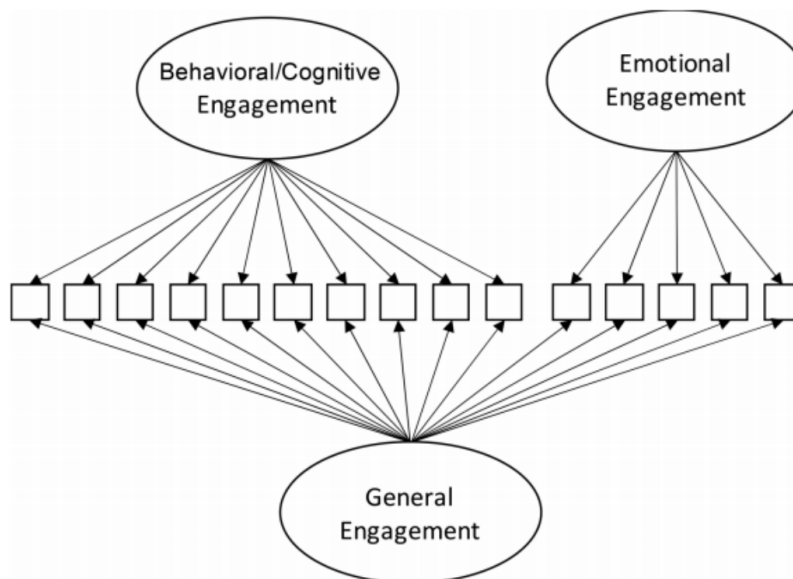


Figure 4. Fifteen-item AcES BC-E bifactor model with an overarching general engagement factor that accounts for commonality between all items, as well as separate behavioral/cognitive and emotional engagement factors that represent engagement for those dimensions above and beyond what is included in the general engagement factor.

图4. BC-E模型（15项AcES）

5. 双因子模型：BC-E模型（15项AcES），引入方法因子。对模型5对应的15项AcES进行项目分析，结果表明所有项目负荷均具有统计学意义（ $p < 0.05$ ）。

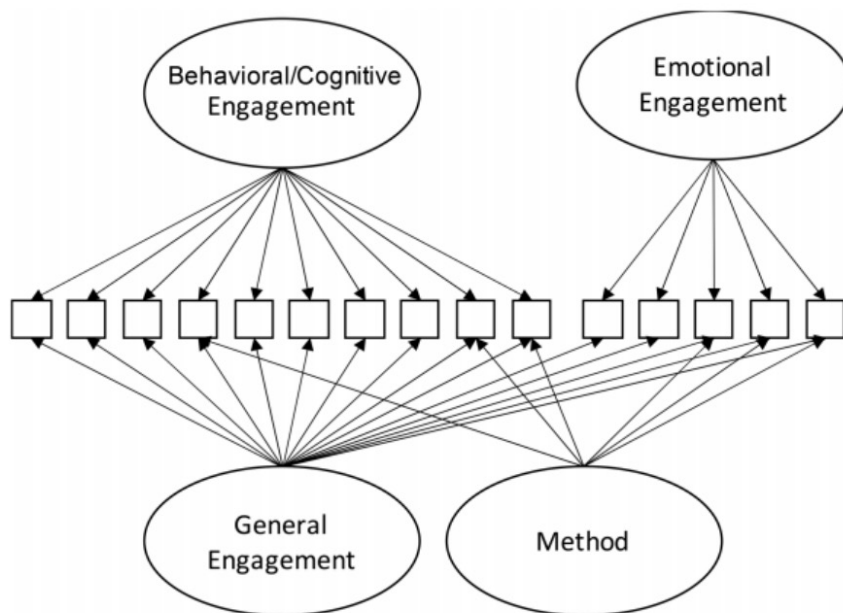


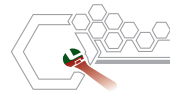
Figure 5. Fifteen-item AcES BC-E bifactor model with a negative method factor that accounts for the relation between negatively worded items.

图5. BC-E模型（15项AcES，引入方法因子）

总而言之，用AcES收集的数据的CFA结果表明，具有负方法因素和结合行为/认知的双因素模型E因素显示了良好的数据模型拟合。这是最受支持的AcES BC-E模型（仅包括行为、认知和情绪项目）以及AcES BC-E-S（包括行为、认知、情感和社会项目）。双因素模型的存在表明，在这些工作表活动中的参与度可以建模为学生在活动中的整体参与度。具有与特定因素相关的单独成分的项目，允许同时测量整体和单独成分。在未来的研究中，双因素模型可以允许进一步同时探索某些学生的学习结果如何与整体参与相关，以及行为、认知、情感和社会参与的单独因素。

该研究对一线教师的实践具有理论意义。教师可以利用该研究所开发的化学通识课程学生活动参与度测量工具（AcES）来收集学生对于不同学习活动的反馈，从而指导教师更有针对性地开展学习活动。

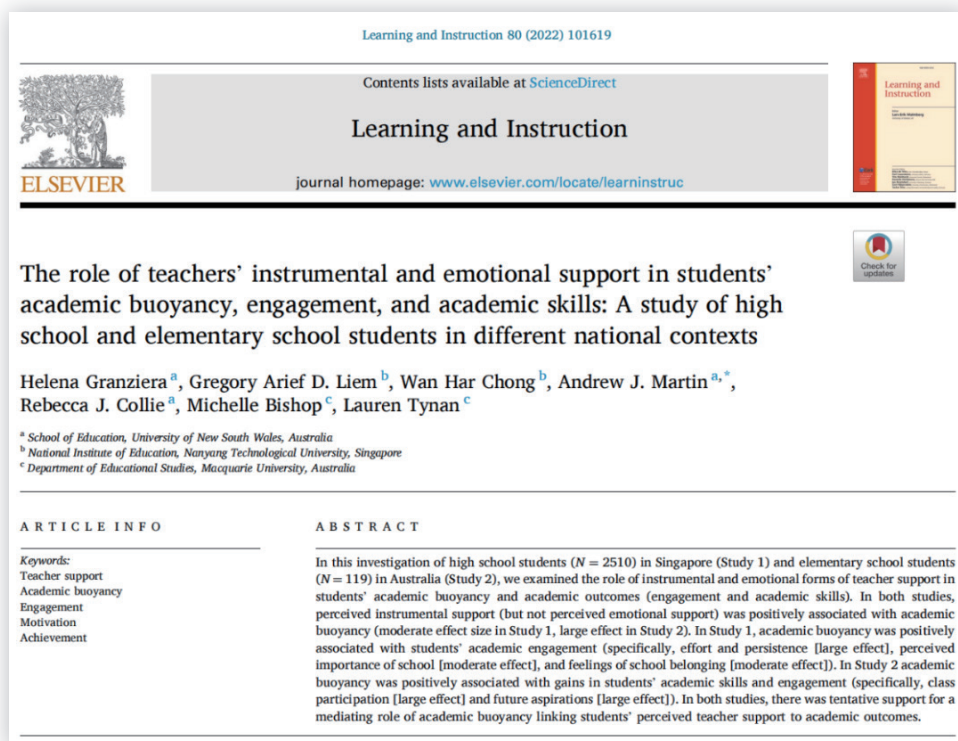
（本文责任编辑：曾思莹；本文责任校对：曾思莹）



《教师工具支持和情感支持在学生学业浮力、参与度和学术技能中的作用：针对不同国家背景下的高中生和小学生的研究》推介

皇甫倩，罗舟颖

(西南大学 教师教育学院)



原文：Helena Granziera, Gregory Arief D. Liem, et al. The role of teachers' instrumental and emotional support in students' academic buoyancy, engagement, and academic skills: A study of high school and elementary school students in different national contexts[J]. *Learning and Instruction*, 2022, 80, 101619.

原文网址：<https://doi.org/10.1016/j.learninstruc.2022.101619>

核心概念：教师支持（teacher support）；学业浮力（academic buoyancy）；参与度（engagement）；动机（motivation）；成就（achievement）

应对学业挑战的能力对学生的学业成就很重要。学业浮力（academic buoyancy）是指克服学校“日常”生活中典型的学业挑战的能力，它已经成为与学生学业成就相关的一个特别重要的属性。例如：学业浮力与高水平的参与度、自尊、动机、自我调节、享受、抱负和学业成就有关，也与低水平的无聊、绝望、情绪不稳定、神经质和焦虑有关。因此，学业浮力显然与许多学术和个人的成就相关。然而，作为一个最近发展起来的观念，促进学业浮力的积极因素研究仍然是有待进一步探索的。



来自澳大利亚的新南威尔士大学（University of New South Wales）、麦考瑞大学（Macquarie University）和新加坡的南洋理工大学（Nanyang Technological University）的7位教育领域的学者合作发表了一篇题为《教师的工具支持和情感支持在学生学业浮力、参与度和学术技能中的作用：针对不同国家背景下的高中生和小学生的研究》的论文，他们对新加坡高中生（N=2510）（研究1）和澳大利亚小学生（N=119）（研究2）进行了调查，主要探究了教师支持（工具性支持和情感性支持）和学业浮力在感知到的教师支持和学生学业成就（参与度和学术技能）之间的中介作用，并构建了理论模型（如图1所示）。该研究利用了两个相互独立收集到的数据集，考虑了事后协同效应，进行整合数据分析。

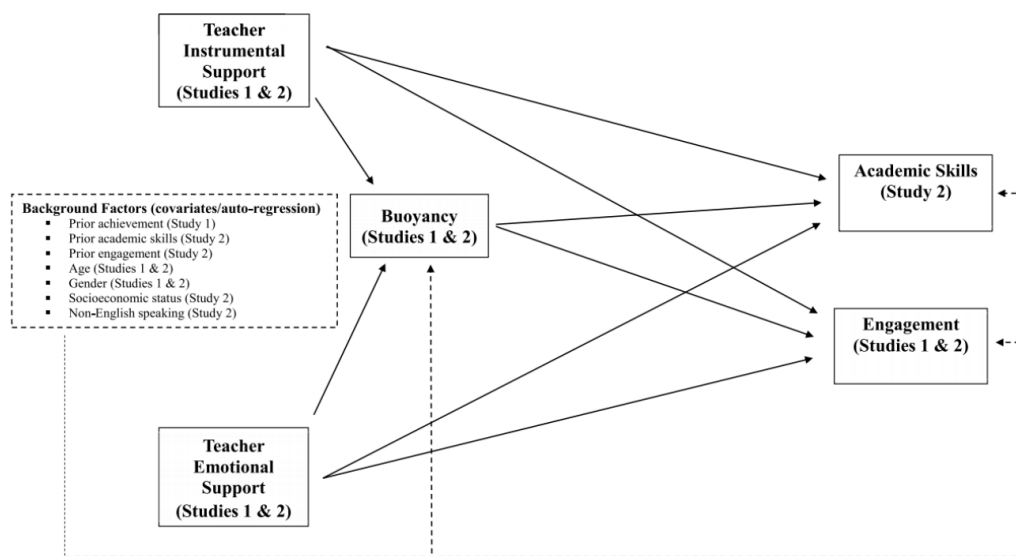
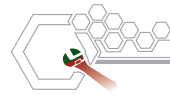


图1. 理论模型：研究1和2

Figure 1. Hypothesized model: studies 1 and 2.

理论模型中所涉及的核心概念界定如下：学业浮力是学校生活过程中典型的低水平或轻微逆境的经历，例如学生可能会面临的考试压力、有截止日期的竞争性作业和成绩不佳等挑战。教师支持划分为工具支持和情感支持两个维度。其中工具支持（instrumental support）指的是教师提供仪器资源和实际帮助来帮助学生解决学术问题和发展学术技能，也被认为是促进行动的支持，例如：提问学生、澄清概念、纠正错误和学习行为建模。而情感支持（emotional support）的特点是关心、照顾和尊重学生，了解学生的思想、感受和行为，温暖学生和以友善的态度对待学生。学业成就（academic outcomes）包括参与度（engagement）和学术技能（academic skills）等。参与度（engagement）由情感、认知和行为维度组成，其中情感参与包括愉悦感、学校归属感；行为参与包括参与、努力和做作业；认知参与包括抱负和对学校重要性的感知。



研究1

这是一项横断面研究。样本是来自新加坡10所中学的2510名学生，其中女生占51.8%，男生占48.2%，8年级的学生占49.3%，九年级的学生占50.7%，学生的平均年龄为14.53岁（SD=0.53）。调查耗时约45分钟，由各学校的教师在课堂环境中进行。测量教师支持的量表是Furman和Buhrmester（1985）开发的关系库存网络的子量表，采用李克特五级评分法，教师工具支持由3个题目来衡量（ $\Omega=0.81$ ）；教师情感支持也由3个题目来衡量（ $\Omega=0.81$ ）。测量学业浮力的量表是学业浮力量表（ABS），由4个题目组成（ $\Omega=0.89$ ），采用李克特七级评分法。该研究中，将参与度视为样本的学业成就，测量参与度的量表是学校态度评估调查自律分量表修订版SAAS-R的6题（ $\Omega=0.92$ ）；测量对学校重要性的感知的量表是目标评估分量表的6题（ $\Omega=0.94$ ）；测量学校归属感的量表是SAAS-R的5题（ $\Omega=0.96$ ），采用李克特七级评分法。此外，同时将性别、年龄、成绩（小学毕业考试总分，含英语、数学、科学和母语）作为协变量。

收集到数据后，分析的初始阶段是使用SPSS v26计算所有变量的平均值、标准差、偏度和峰度，用于评估分量表的描述性和分布特征。在此之后，采用验证性因子分析（CFA）来评估假设模型中包含的潜变量的结构效度。并使用由该CFA产生的项目因子载荷和误差方差来估计系数 ω ，以此作为子量表信度的指标。同时CFA还提供了所有潜变量之间的相关性。分析的主要阶段是在结构方程建模（SEM）框架内进行的，以估计各变量之间的假设关系，此外，还评估了该模型中的中介效应是否显著。

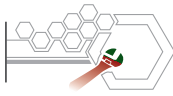
描述性统计分析结果表明，6个子量表的得分大致呈正态分布，偏度和峰度的值在可接受范围内。此外，量表具有较高的信度（0.81~ ω ~0.96）和效度。

CFA的结果显示了所有变量之间的潜在相关性（如表1所示）：（1）教师工具支持与教师情感支持、学术浮力都与参与度的三个维度呈正相关；（2）参与度的三个维度之间两两呈正相关；（3）协变量方面，年龄与学校归属感负相关、性别与学校归属感相关，其中样本中的男孩在学业浮力和学校归属感方面高于女孩。

表1. 研究1：变量之间的潜在相关性

Table 1. Study1: Latent correlations among factors.

	1	2	3	4	5	6	7	8	9
1. Age	-								
2. Gender (Male)	.01	-							
3. Prior Achievement	-.18	-.15*	-						
4. Teacher Instrumental Support	.05	.07	-.05	-					
5. Teacher Emotional Support	.01	.01	-.06	.77***	-				
6. Academic Buoyancy	.01	.09**	-.03	.28***	.23***	-			
7. Feelings of School Belonging	-.08***	.15***	.04	.40***	.31***	.33***	-		
8. Effort and Persistence on Schoolwork	-.05	.04	.02	.35***	.31***	.40***	.42***	-	
9. Perceived Importance of School	-.06**	.03	.03	.21***	.25***	.14***	.30***	.48***	-



采用SEM对图1展示的假设模型进行了检验，结果表明该模型产生了极好的拟合， χ^2 (372, N=2510) = 2124.08, $p < 0.001$, CFI=0.954, RMSEA=0.043。表2报告了所有标准回归系数，图2显示了变量之间的相关性（只显示 $\beta \geq 0.10$ ）。结果表明：教师工具支持与学业浮力呈正相关（ $\beta=0.22$, $p < 0.001$, 中等效应），学业浮力与努力、做作业呈正相关（ $\beta=0.33$, $p < 0.001$, 大效应），与对学校重要性的感知呈正相关（ $\beta=0.15$, $p < 0.001$, 中等效应），与学校归属感呈正相关（ $\beta=0.24$, $p < 0.001$, 中等效应）；教师工具支持与努力、做作业直接相关（ $\beta=0.21$, $p < 0.001$, 中等效应），与对学校重要性的感知直接相关（ $\beta=0.25$, $p < 0.001$, 大效应），与学校归属感直接相关（ $\beta=0.29$, $p < 0.001$, 大效应）；教师情感支持与学业浮力无关，然而，它与努力、做作业直接相关（ $\beta=0.13$, $p < 0.001$, 中等效应）。对中介效应（教师支持→学业浮力→参与度）的检验显示了所有六种可能的中介效应的统计结果显著（ $p < 0.05$ ）：教师工具支持→学业浮力→学校归属感/对学校重要性的感知/努力做作业（ $\beta=0.05/.03/.07$, $p < 0.001$ ）和教师情感支持→学业浮力→学校归属感/对学校重要性的感知/努力、做作业（ $\beta=0.02/.01/.03$, $p < 0.05$ ）。最显著的中介效应出现在教师工具支持→学业浮力→努力、做作业（ $\beta=0.07$, $p < 0.001$ ）的路径中。

表2. 研究1：结构方程模型中的非标准化和标准化回归系数及其对应的标准误差

Table 2. Study 1: Unstandardized and standardized regression coefficients and their corresponding standard errors from structural equation model.

	Academic Buoyancy			Feelings of School Belonging			Perceived Importance of School			Effort and Persistence on Schoolwork		
	B(SE)	B/SE	β	B	B/SE	β	B	B/SE	β	B	B/SE	β
Teacher Instrumental Support	.32(.05)	6.99	.22***	.49(.03)	16.05	.29***	.23(.03)	6.52	.25***	.17(.04)	4.06	.21***
Teacher Emotional Support	.13(.06)	2.38	.09**	.13(.06)	2.37	.08*	-.05(.04)	-1.14	-.05	.28(.03)	9.69	.13***
Academic Buoyancy	-	-	-	.27(.03)	9.52	.24***	.09(.02)	5.73	.15***	.30(.03)	11.89	.33***
Age	.01(.05)	.10	.01	-.21(.04)	-5.49	-.09***	-.09(.03)	-3.18	-.07**	-.11(.05)	-2.22	-.06*
Gender (Male)	.18(.08)	2.31	.08*	.33(.08)	4.29	.13***	-.01(.04)	-.01	-.01	.01(.06)	-.12	.01
Prior Achievement	.01(.01)	-.28	-.01	.01(.01)	2.06	.08*	.01(.01)	.98	-.03	.01(.01)	1.33	.04

Note. Age (in years), Gender (0 = female, 1 = male); * $p < .05$, ** $p < .01$, *** $p < .001$; B = Unstandardized Regression Coefficient, SE = Standard Error, β = Standardized Regression Coefficient, B/SE = Unstandardized Regression Coefficient/Standard Error.

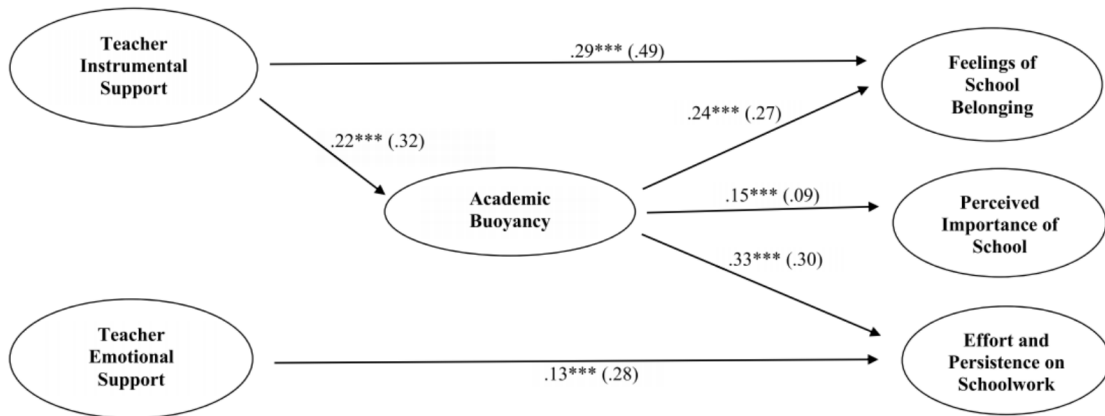
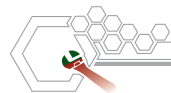


图2. 研究1：结构方程模型中的显著影响：标准化回归系数

Figure 2. Study 1: Significant effect from structural equation model: standardized beta coefficients.



研究2:

这是一项纵向研究，它在不同的国家背景下拓宽了研究1。样本是来自澳大利亚悉尼五所小学的119名学生，时间1（Time-1）的学生平均年龄为10.9岁（SD = 0.73）。样本中超过一半（54.2%）为女生，地区社会经济指数平均值为1038.64（SD = 11.41），略高于全国平均水平1000（SD = 100）。其中有18.2%的样本被确定为有特殊需要（阅读障碍、计算障碍或书写障碍），29.2%的学生表示他们在家中说英语以外的语言。参与者在课堂时间（或另一个最适合特定班级参与的指定校内时间）完成了在线问卷以及算术和识字测试。在Time-1问卷和测试后6-9个月进行Time-2问卷和测试，两个测试题目相同。

测量教师支持的量表是师生关系量表-学生（TSRI-S）的2个分量表：教师工具支持由5个题目来衡量（ $\Omega=0.81$ ）；教师情感支持也由5个题目来衡量（ $\Omega=0.92$ ），采用李克特七级评分法。测量学业浮力的量表同研究1一样，均是ABS，由4个题目组成（ $\Omega=0.66$ ）。测量参与度与研究1所提出的三维度一致，而使用的量表是Martin和Marsh开发的量表：参与功课的由4个题目来衡量（ $\Omega=0.88$ ）；抱负也由4个题目来衡量（ $\Omega=0.78$ ）；愉悦感同样由4个题目来衡量（ $\Omega=0.90$ ）。除了研究1所提出的参与度之外，增加了学术技能，二者共同视为学业成就。测量学术技能的量表包括简短的识字和算术测验（ $\Omega=0.89$ ），共10题，难度逐渐增加。识字测验改编自Green等人验证的测量方法，是对拼写，速度和理解的综合衡量；算术测验改编自Martin等人验证的较长测验，包括对加法、减法、除法、乘法、分数、小数、百分比和基本代数进行评估的题目。此外，将性别变量、是否有特殊需要、惯用语言是否为英语、年龄、社会经济地位等视为协变量。

参与度由情感、认知和行为维度组成，其中情感参与包括愉悦感、学校归属感；行为参与包括参与、努力和做作业；认知参与包括抱负和对学校重要性的感知。

数据分析方法与研究1类似，但是由于样本量相对较小，Time-1和Time-2的数据分析都采用了路径分析。而且为了解释测量误差，还计算了所有实际变量（即除背景变量外的所有变量）的校正误差的平均值。

描述性统计结果表明，7个子量表的得分近似呈正态分布，因为偏度和峰度的值在可接受范围内。子量表的以 ω 系数为代表的信度范围为0.66到0.92，这表明有足够的信度。此外，量表具有良好的信度（ $0.66 \sim \omega \sim 0.92$ ）和效度。

表3显示了所有变量之间的相关性，结果表明：教师工具支持与教师情感支持、愉悦感、抱负呈正相关；教师的情感支持与愉悦感、抱负呈正相关；学业浮力与愉悦感、抱负、参与度呈正相关；抱负与愉悦感呈正相关；参与度与抱负、愉悦感呈正相关；学术技能与参与度呈正相关；在协变量方面，没有发现年龄、性别或非英语背景与任何实质性变



量之间的相关性，社会经济地位与教师情感支持呈正相关，而特殊需求与学术技能呈负相关。

表3. 研究2：变量之间的相关性
Table 3. Study 2:Correlations among factors.

	1	2	3	4	5	6	7	8	9	10	11
1. Age	-										
2. Gender (Male)	.05	-									
3. SES	-.20	-.01	-								
4. Non-English Speaking	.08	-.11	-.11	-							
5. Special Needs	.11	-.04	-.07	.14	-						
6. Teacher Instrumental Support	-.18	.12	.11	.05	.10	-					
7. Teacher Emotional Support	-.08	.10	.16**	.14	-.06	.79***	-				
8. Academic Buoyancy	-.01	-.01	-.15	.12	-.07	.27	.15	-			
9. Enjoyment	-.07	.15	-.07	.02	-.12	.48***	.53***	.28*	-		
10. Future Aspirations	.06	-.04	-.01	.05	-.12	.29*	.30**	.53***	.66***	-	
11. Participation	-.13	-.10	.06	-.01	-.14	.15	.19	.43***	.44**	.79***	-
12. Academic Skills	.14	-.11	-.11	.02	-.32**	-.11	.05	.21	.06	.29	.32*

Note. All correlations shown reflect data collected at Time 2. Correlation between Enjoyment and prior Enjoyment, $r = .46, p < .001$; Correlation between Future Aspirations and prior Future Aspirations $r = .71, p < .001$; Correlation between Participation and prior Participation, $r = .63, p < .001$; Correlation between Academic Skills and prior Academic Skills, $r = .53, p < .001$; SES = socio-economic status; All substantive variables (instrumental support, emotional support, buoyancy, enjoyment, future aspirations, participation, academic skills) are error-adjusted.
* $p < .05$, ** $p < .01$, *** $p < .001$.

图3显示了实质性变量之间的统计学上的显著关系。结果表明：模型得到了很好的拟合， $\chi^2(20, 21.27)$ ， $p=0.168$ ， $RMSEA=0.053$ ， $CFI=0.981$ 。教师工具支持与学业浮力呈正相关（ $\beta = 0.42, p < 0.05$ ，大效应）；学业浮力与学术技能（ $\beta = 0.23, p < 0.05$ ，中等效应）、参与度（ $\beta = 0.27, p < 0.05$ ，大效应）和抱负呈正相关（ $\beta = 0.33, p < 0.001$ ，大效应）；教师情感支持与愉悦感呈正相关（ $\beta = 0.49, p < 0.01$ ，大效应）；然而，教师情感支持与学业浮力、参与度、抱负或学术技能无关。对于中介效应（教师支持→学业浮力→学术技能）的后续分析显示没有统计学意义上的关系（ $p < 0.05$ ）。

表4. 研究2：路径分析中的非标准化和标准化回归系数及其相应的标准误差
Table 4. Study 2:Unstandardized and standardized regression coefficients and their correspondings standard errors from path analysis.

	Academic Buoyancy			Enjoyment			Future Aspirations			Participation			Academic Skills		
	B (SE)	B/SE	β	B (SE)	B/SE	β	B (SE)	B/SE	β	B	B/SE	β	B (SE)	B/SE	β
Teacher Instrumental Support	.29 (.16)	1.86	.42*	-.04 (.17)	-0.21	-.03	.03 (.12)	0.25	.05	-.07 (.18)	-0.40	-.09	-.23 (.27)	-0.86	-.17
Teacher Emotional Support	-.13 (.16)	-.79	-.18	.50 (.16)	3.10	.49**	.14 (.11)	1.24	.22	.13 (.12)	1.10	.18	.20 (.28)	0.73	.17
Academic Buoyancy	-			.08 (.15)	0.54	.08	.32 (.10)	3.12	.33***	.31 (.10)	3.12	.27*	.45 (.21)	2.17	.23*
Prior Enjoyment	-			.48 (.12)	3.86	.49***	-			-			-		
Prior Future Aspirations	-			-			.55 (.09)	6.03	.67***	-			-		
Prior Participation	-			-			-			.63 (.10)	6.13	.59***	-		
Prior Academic Skills	-			-			-			-			.49 (.09)	5.15	.48***

Note. All coefficients shown reflect data collected at Time 2, with the exception of those indicated as 'prior'. All substantive variables (instrumental support, emotional support, buoyancy, enjoyment, future aspirations, participation, academic skills) are error-adjusted; For parsimony, covariates that did not attain statistical significance ($p < .05$) were removed successively (starting with lowest standardized parameter) and new iterations of the model were run: this process revealed no covariates to retain in the model.

B = Unstandardized Regression Coefficient, SE = Standard Error, β = Standardized Regression Coefficient, B/SE = Unstandardized Regression Coefficient/Standard Error.

* $p < .05$, ** $p < .01$, *** $p < .001$.

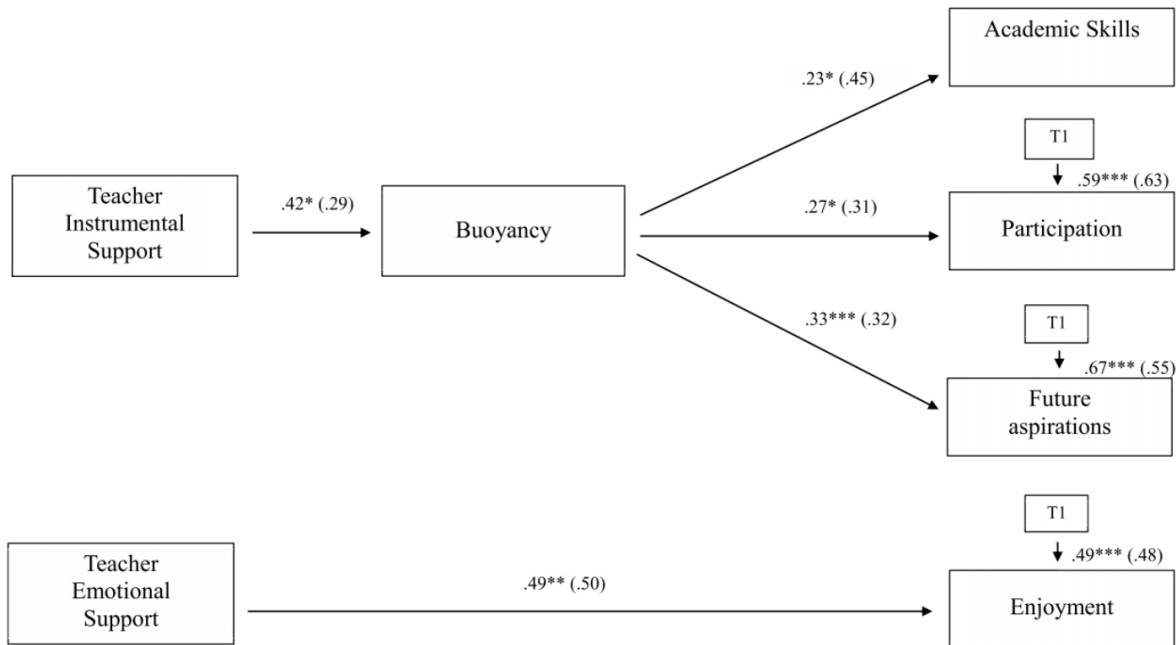
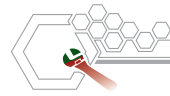


图3. 研究2: 路径分析中的显著影响: 标准化回归系数

Figure 3. Study 2: Significant effects from path model: standardized beta coefficients.

综上所述, 在这两项研究中, 感知到的教师工具支持与学业浮力都呈正相关(研究1中的效应中等, 研究2中的效应较大)。在研究1中, 学生的学业浮力与参与度呈正相关(特别是努力和做作业(大效应)、对学校重要性的感知(中等效应)和学校归属感(中等效应)); 在研究2中, 学生的学业浮力与学术技能和参与度(特别是参与度(大效应)和抱负(大效应))呈正相关。此外, 在研究1中, 教师工具支持与努力、做作业(中等效应)、对学校重要性的感知(大效应)、学校归属感(大效应)直接相关; 情感支持与学业浮力无关、与努力和做作业直接相关(中等效应)。并且, 在研究2中, 教师情感支持与愉悦感呈正相关(大效应); 然而, 教师的情感支持与参与度、抱负或学术技能无关。这两项研究初步支持了学业浮力在学生感知到的教师支持与学业成就之间起到中介作用的理论假设。

最后, 该研究具有一定的实践意义。首先, 教师可以通过提供工具性或情感性不同形式的支持来促进学生的学业成就。其次, 提高教师对如何有效实施工具支持和情感支持的认识。最后, 通过学业浮力来增强学生的参与度。

(本文责任编辑: 曾思莹; 本文责任校对: 曾思莹)

华东师范大学《学科教育研究》 内部交流稿件征集启事

各位老师：

《学科教育研究》是华东师范大学主管和主办的上海市连续性内部资料性出版物，于2021年5月创办，主要介绍国内外学科教育研究和教师教育研究的最新成果，出版物仅限华东师范大学内部交流。

《学科教育研究》目前以编发国内外学科教育和教师教育最新成果的推介类文章为主，同时编发学科教育领域的研究综述、中小学名师成长历程、学校学科教育和教师教育工作进展等方面稿件。如果您在研究过程中了解到一篇高质量的学科教育研究的前沿成果，就可以以文章推介或评论的形式撰文推荐，如果你全面了解了学科教育领域某些研究方向的一系列成果和动态，则可以将研究综述投给我们，从而为本校的学科教育和教师教育的研究者提供参考。您也可以指引或带领您的学生共同完成这些工作，以促进学生对相关领域和成果的了解。我们希望您更多地推介来源于SSCI和CSSCI期刊的最新成果，推介类文章要能详细地介绍作者信息、研究方法、样本范围、结果图表和研究结论等内容，长度要求不少于3000字（含图像和排版空白所占版面），并要求提供原文章的DOI号、访问网址和原文基本信息截屏。

《学科教育研究》暂设语文教育、外语教育、数学教育、物理教育、化学教育、生物教育、科学教育、思想政治和公民教育、历史教育、地理教育、教师教育等栏目，也欢迎体育与健康、艺术、通用技术、信息技术等其他学科教育方向的成果推介稿件。希望《学科教育研究》能为华东师范大学从事学科教育和教师教育的研究者提供一个了解国内外前沿成果和最新动态的平台，以促进我校学科教育和教师教育研究水平的进一步提高，这也是《学科教育研究》这本连续性内部资料性出版物创办的宗旨和意义。

《学科教育研究》由华东师范大学教师教育学院承办，上海华教印务有限公司承印，大16K开本、每期48页、双月刊，准印证号（K）0904。根据《内部资料性出版物管理办法》（国家新闻出版广电总局令2015年第2号）和《上海市新闻出版局关于开展连续性内部资料自查自纠及专项整治工作的通知》（沪新出版[2018] 80号）的相关规定，《学科教育研究》不能发表学术论文，出版物只向校内员工发放，不向外单位发放。稿件一般来自华东师范大学的教职工（或教师携学生共同完成），我们也将通过约请的方式欢迎校外的老师提供学科教育研究的信息供我校教师参考。

欢迎华东师范大学学科教育研究方向的教师为《学科教育研究》提供稿件，也欢迎您指导学生完成相关文献的深度阅读和撰写推介或综述的工作，感谢您对这份出版物的关注和支持！

咨询邮箱：aprilgreen@yeah.net

稿件接收平台：<http://xkjyyj.paperonce.org/#/>

《学科教育研究》编辑部

采编员招聘

《学科教育研究》在华东师范大学教师教育学院范围内招聘硕士研究生担任各专业采编员，主要负责本专业范围内的稿件组稿、编辑、校对等工作，报酬根据组稿后录用稿件的编辑校对文稿量计算，标准为实际排版页每页30元，欢迎细心、认真、负责的同学报名。报名平台为：<https://www.wjx.cn/vm/wVvAmKl.aspx>。可长期随时通过本平台报名，报名后联系编辑部（aprilgreen@yeah.net）进行面谈交流，以确认是否聘用。

《学科教育研究》编辑部

《学科教育研究》推介稿投稿格式规范和自查表 (下载最新版)

投稿示例文章 (另提供)

(2023年第1期起执行, 2023.02.10版)

投稿格式提示图 (另提供)

序号	类别	投稿格式要求	自查标记
1.	推介内容 (重视)	总结、介绍英文S刊上的论文,应在理解意思后表意,不是直译。首先要 自己完全读懂原文 ,推介要用 中文说得通的语言、要用自己明白意思的句子 表达原文意思	
2.		要尽可能展示或概括总结原文的研究目的、思路、方法、结果、结论,包括样本是如何选择的,数据是如何采集的,欢迎使用原文中的图和表(欢迎中英对照)	
3.	核心概念确定 (重要)	文中出现的 核心概念 或 关键词 译称要全文统一,全文均使用概念术语的 中文全称 可以与原文关键词不同。是作者需要统一和强调的概念	
4.	字体字号	为简化格式要求,全文(包括标题)均为宋体、五号,表格标题和内容可以用小五号	
5.	格式和排版	保持常态、美观,不作其他要求(排版格式将由编辑部统一)	
6.	标题	为“《...原文译名...》推介”,原文译名与正文中提到的原文译名完全一致	
7.	作者署名	多位作者之间用“,”,二字作者中间加一空格,如“刘 兵, 张小鹏”	
8.		作者如为学生,作者中需要有指导教师,推介需要得到教师指导并认可	
9.	作者单位	署在作者下一行,高校名称和二级院系之间有一半角空格,中学和教研室等不加二级单位,括号为中文括号。如多位作者来自不同的单位,需要在作者署名处用上标表示不同单位,作者单位处用1.....; 2.....标记单位,中间用中文分号隔开,如: “(1.华东师范大学 教师教育学院; 2.南京市教育局教研室)”	
10.	原文截图	可对原始截图中的内容作剪除和位置调整,图片加阴影	
11.	原文文献信息	形式参考如下样例,由作者、原文标题、出版信息(期刊、年卷期页)三部分组成 原文: Tothova M, Rusek M, et al. Students' procedure when solving problem tasks based on the periodic table: An eye-tracking study[J]. Journal of Chemical Education, 2021, 98(6): 1831-1840.	
12.		作者姓和名首字母大写	
13.		作者中间用半角“,”,超过2位作者的,第三位及以后的作者用“,et al”替代	
14.		标题和副标题的首词的首字母和专有名词的首字母大写,其他单词字母均小写	
15.		期刊名称所有实词和大于等于5个字母的虚词首字母大写	
16.		年卷期页的格式如“2021, 120(5): 165-175.”中间为半角符号,有些期刊格式可特殊	
17.	原文网址	以“原文网址: https://”开头,加超链接,可点击其打开文章页面	
18.	核心概念	核心概念中英文对照,括号用中文括号。如:“教育公平(educational equity)”	
19.		英文除专有名词首字母大写外,一般单词字母均小写	
20.		概念之间用中文分号隔开,最后一个核心概念后面没有标点符号	
21.	图表	图表均“嵌入式”或“无环绕”排在正文中。图表名就是正常(居中)的正文文本	
22.		每个图表均需要有图表名称,表名在表上,图名在图下,图表名须中英对照	
23.		图表序号和图表名称中间加一半角空格,如“图1. 教师...”、“表1.”,英文图、表两个单词用“Figure 1. Te...”“Table 1.”;英文图表名最后一般有“.”号	
24.		英文图表名首字母大写,其他均小写,如“Figure 1. The proposed moderated.....”	
25.		图表内容也尽量中英文对照(根据版面情况和美观适宜性综合判断,不作绝对要求)	
26.	正文	所推介的文章要说出原文标题(与推介标题中的原文译名一致)	
27.		要介绍作者姓名及其机构	
28.		文中涉及的机构和地名中英文对照,英文加括号,不知名的机构和地名直接用英文	
29.		所推介的文章要用“原文”指代,不能用“本文”	
30.		缩写概念在正文首次出现时要中文括注英文,如:兴趣区(areas of interest, AOI)	
31.	正文中的评述	推介正文最后要有一段对原文的意义、问题的评价(可以来自原文作者,更希望推介作者的评论),并叙述对读者的启示(今后进一步的研究方向、实践指导等)	
32.	全文长度	推介全文尽量保持4-6页(按期刊的字号和行距估计)	
33.	参考文献	不需要参考文献,如有则应删除	
34.	原文PDF	投稿时同时提供原文PDF	
35.	本自查表	打印纸质表,投稿时同时提供各项均合格的本自查表拍照版(右侧做合格标记)	

在使用本表前请务必下载最新版

作者签名:



期刊网站

《学科教育研究》编委会

主 编：代蕊华 朱 梅
副主编：汪晓勤 华春燕 蒋 瑾

《学科教育研究》编辑部

本刊责任编辑：卢晓旭
本期责任编辑：曾思荧 安爱玲
咨询邮箱：aprilgreen@yeah.net
稿件接收平台：<http://xkjyyj.paperonce.org/#/>

学科教育研究

(上海市连续性内部资料性出版物)
主管：华东师范大学
主办：华东师范大学
承办：教师教育学院
承印：上海华教印务有限公司
印数：300册
准印证号：(K)0904
发行对象：华东师范大学学科教育研究机构师生
2023年 第2期 2023年3月1日出版
(内部资料，免费交流)
