

## Journal of Multitiered System of Supports

# 初中普遍性社会情感学习项目的实施与学生发展成果

Lauren Evanovich\*, 南佛罗里达大学  
Nicholas A. Gage, 南卡罗来纳大学  
Robin Parks Ennis, 阿拉巴马大学  
Tracy Ulrich, WestEd

通信作者: Lauren L. Evanovich, PhD, 4202 E Fowler Avenue, EDU105, Tampa, FL 33620; [evanovich@usf.edu](mailto:evanovich@usf.edu).

### 摘要

本研究历时两年，评估了Second Step Digital Middle School Program这一普遍性社会情感学习（SEL）项目对美国东南部某大型学区4,903名初中生的影响。研究采用准实验方法，发现高忠实度实施Second Step项目可使英语语言艺术（ELA）成绩获得幅度较小但具有统计学显著性的提升，大幅减少违规行为（包括转介至学校管理办公室处理和停学处分），并提高出勤率（每年约增加2.5天）。此外，实施Second Step项目学校的学生报告了更好的师生关系、更强的学校归属感，以及对学校氛围更积极的评价。然而，学生的自我管理能力和同伴支持关系未出现统计学显著变化。研究结果凸显了在充满挑战的初中阶段持续开展SEL项目对于促进学生多维度发展的价值，同时指出了缺乏个体层面学业基线数据及结果可推广性等方面的局限。

关键词：社会情感学习、Second Step、学校氛围、初中、学生发展成果

初中是学生发展的关键时期，其身份认同、同伴关系和学业要求均经历深刻变化。在这一阶段，学生常出现学习动机下降、行为困难增多、学业表现下滑及学校联结感减弱等情况。美国全国性数据反映出令人担忧的趋势：青春期早期纪律事件增加，出勤率下降，学生在初中阶段报告的情绪幸福感也随之降低。这些挑战凸显了学校有必要采取能够同时改善学业投入、行为表现和社会情感健康的策略。本研究考察普遍性社会情感学习（SEL）项目Second Step对学生学业成就、行为、出勤及社会情感幸福感的影响。

### 初中阶段的挑战

从发展角度看，初中阶段的身份认同、同伴互动和学业期望均发生显著转变。青少年在接触新的社会环境和不断变化的同伴群体时，开始重新审视并建构自我概念（Onetti et al., 2019）。既有研究表明，积极参与学校生活不仅有助于学业进步，也支持连贯身份认同的形成；学校倦怠则可能阻碍这一过程（Erentaitė et al., 2018）。此外，随着青少年日益寻求自我界定，族裔认同和社会认同等因素更加突出，进一步体现了个体发展与学校环境之间复杂的相互作用（Erentaitė et al., 2018）。

升入初中往往意味着更高的学业要求和新的组织结构，这些变化可能与青少年日益增长的自主需求发生冲突。当学业期望提高却缺乏相应的教师支持时，这种不匹配可能导致内在动机下降。实证证据显示，当学生认为教师具有支持性——既保持高期望，又实施个性化教学——其动机水平往往更高（Hornstra et al., 2018）。相反，缺乏这些支持性做法可能加剧动机下降，并在过渡阶段带来不良行为后果。此外，从小学到初中的过渡还与学业表现下降、行为困难增加等一系列负面结果有关，这在很大程度上源于适应新教育环境的压力（Akos, 2006）。

初中阶段的另一个关键问题是学校联结感的减弱。实证研究发现，学生对学校共同体的归属感和支持感受到干扰，与学业成就下降及纪律问题增加密切相关（Niehaus et al., 2012）。随着学业压力上升和同伴关系变化，联结感较弱的学生可能更容易出现行为问题，并疏离学校活动，从而进一步加重学业困难（Niehaus et al., 2012）。综合来看，初中是一个复杂的发展阶段，动机下降、新出现的行为困难和学校联结感减弱，共同对学业表现与个人发展构成挑战。

普遍性SEL课程因有望促进学生的人际技能、自我调节和心理韧性而受到关注。通过主动将SEL教学融入日常学校生活，这些项目旨在为改善行为、增强学校归属感及提高学业表现奠定基础。尤其重要的是，普遍性SEL面向所有学生，而不仅限于被识别为存在风险的学生，同时促进技能培养机会和预防性支持的公平可及。

越来越多的研究证明，SEL项目对多种学生发展成果具有积极作用。元分析表明，SEL干预与学生情绪调节、社会能力和学校态度的改善，以及品行问题和纪律转介的减少有关（Durlak et al., 2022）。此外，参与SEL还与幅度较小但具有实际意义的学业成就提升相关。不过，效果的大小和一致性有所差异，尤其取决于发展阶段与实施情境。青春期早期学生自主性增强，同伴关系也更为复杂，使初中SEL实施面临独特挑战。

Second Step Middle School Program是美国学校使用最广泛的普遍性SEL课程之一，通常在多层次支持系统的第一层级实施。该项目以认知行为理论和社会情感发展理论为基础，着重教授情绪调节、共情、问题解决和负责任的决策。既有评估表明，Second Step可减少欺凌与攻击行为、增加亲社会行为，但相对较少的研究在严谨分析框架和真实学校情境下，考察其对出勤、学业表现等更广泛学校功能指标的影响。

## 研究目的

针对初中普遍性SEL项目的评估，很少同时全面考察学生多个功能领域。现有文献大多单独关注行为或社会情感结果，而未考虑这些维度如何与学业投入和学校氛围交织。本领域日益需要整合多个学生层面结果、采用纵向设计并在真实学校情境中评估效果的研究。

本研究通过评估Second Step Middle School Program对学业成就、行为、出勤和幸福感的影響，回应上述不足。我们采用控制学校和学生层面因素的双重差分设计，考察接触Second Step是否与课程成绩、纪律结果、出勤，以及学生自报的学校氛围和社会情感技能变化相关。研究旨在严谨、全面地评估Second Step在初中情境中的影响，为学生关键发展时期的普遍性SEL项目提供证据。

## 方法

我们收集了南卡罗来纳州某大型学区四所初中共4,903名学生在2022–2023和2023–2024学年的去标识化个体数据。儿童委员会（Committee for Children）和学区将其中两所学校认定为数字版Second Step项目的高实施学校，即完成超过60%的课程。这两所学校构成干预组。另两所学校虽可使用Second Step，但未完成任何课程，或仅零星完成少量课程，构成比较组。

## 样本

表1列出了学生的描述性统计。两组中白人学生均超过半数。比较组的西班牙裔学生较多，干预组的黑人和多种族学生较多。干预学校中残障学生也更多。

表 1  
分类变量的描述性统计

| 变量             | 比较组   |       | 干预    |       |
|----------------|-------|-------|-------|-------|
|                | 频数    | 百分比   | 频数    | 百分比   |
| 性别             |       |       |       |       |
| 女性             | 3,146 | 49.09 | 2,602 | 47.06 |
| 男性             | 3,182 | 49.65 | 2,896 | 52.38 |
| 非二元性别          | 1     | 0.02  | 1     | 0.02  |
| 缺失             | 80    | 1.25  | 30    | 0.54  |
| 种族             |       |       |       |       |
| 白人             | 3,527 | 55.03 | 3,306 | 59.79 |
| 黑人             | 1,232 | 19.22 | 1,476 | 26.7  |
| 西班牙裔           | 1,227 | 19.14 | 339   | 6.13  |
| 亚裔             | 134   | 2.09  | 90    | 1.63  |
| 美洲原住民          | 9     | 0.14  | 8     | 0.14  |
| 夏威夷原住民或其他太平洋岛民 | 14    | 0.22  | 24    | 0.43  |
| 多种族            | 186   | 2.90  | 256   | 4.63  |
| 缺失             | 80    | 1.25  | 30    | 0.54  |
| 残障状况           |       |       |       |       |
| 无残障学生          | 5,529 | 86.27 | 4,632 | 83.78 |
| 残障学生           | 880   | 13.73 | 897   | 16.22 |

测量

本研究使用学业和行为结果测量收集数据，具体如下。

学业表现

学生的英语语言艺术（ELA）和数学成就以学区报告的期末课程成绩衡量。成绩范围为0至100分，反映整个学年的累计表现。我们将其作为连续结果变量，分数越高表示学业表现越好。这些成绩覆盖多种课程，并在学校和年级之间采用统一标准。

在ELA方面，学生修读不同年级及加速课程，包括普通课程和荣誉课程。常见课程包括六、七、八年级ELA及相应荣誉课程。其他课程还包括英语、英语1荣誉课程，以及阅读辅助和语言艺术拓展等专门课程。

数学情况类似，数据涵盖不同年级和加速课程，包括普通及荣誉课程。这些课程包括普通教育数学课程（如六、七年级数学）、荣誉课程，以及代数1、代数1荣誉课程、几何荣誉课程和多种加速学习路径等进阶课程。

## 行为结果

行为结果来自学区学生信息系统的行政记录。我们采用三个累计指标量化学生整个学年的行为，均作为计数结果变量，数值越高表示违规行为越频繁。

纪律办公室转介（ODR）。学校工作人员因违反校规或品行问题而提交并记录的行为转介总次数。

校内停课处分（ISS）。学生因纪律原因被安排接受校内停课处分的累计次数。

校外停学处分（OSS）。学生被暂停到校且不得进入校园的累计次数。

出勤。学生出勤以学区出勤追踪系统报告的学年到校总天数衡量。该指标反映学生被记录为到校接受教学的天数，并作为学校投入的连续指标。到校天数按计数变量处理，数值越高表示获得教学和参与学习的机会越多。

社会情感幸福感。我们使用Panorama学生调查（Panorama Education, n.d.）。该工具经过验证，旨在了解学生对教学、学习和学校氛围等教育经历的感受，涉及教学效能、学校氛围和学生投入等多个主题。我们采用学区初中实际使用的以下分量表。

师生关系。评估课堂内外教师与学生社会联结的强度。示例题项为：“如果你心情不好地走进教室，老师会多么关心你？”所有题项采用5点李克特量表。

学校归属感。评估学生感到受到学校共同体重视和接纳的程度。示例题项为：“总体而言，你在多大程度上觉得自己属于这所学校？”所有题项采用7点李克特量表。

学校氛围。衡量学生对学校整体社会和学习环境的感受。学区仅使用一个题项：“老师看起来对给你们班上课感到兴奋的频率有多高？”该题项采用5点李克特量表。

自我管理。评估学生不同情境中管理情绪、想法和行为的能力。示例题项为：“过去两周，你做好准备来上课的频率有多高？”所有题项采用5点李克特量表。

支持性关系。考察学生在学校情境中获得支持的关系质量。所有题项均为二分题项。

表2利用本研究收集的数据，列出了Panorama调查每次秋季和春季施测的描述性统计及信度（即内部一致性）。数据呈现出若干重要模式。首先，两个学年秋季得分均略高于春季。其次，除学校氛围仅有一个题项外，所有测量均显示出足够信度（ $>.70$ ）；因此，支持性关系领域的合成信度并不适用。支持性关系领域信度较低，提示应谨慎解释其合成得分结果。表3按领域和测量时间呈现描述性统计，表4则列出相应的原始标准化均值差。

## 数据分析

我们采用基于倾向得分的逆处理概率加权（IPTW），以应对潜在选择偏差，并建立干预组与比较组的基线等值性。倾向得分表示在给定观测特征的条件下，学生被分配到干预组的概率，通过逻辑回归估计。预测变量包括性别、种族/族裔、英语学习者身份、残障学生身份、年级及2022年秋季Panorama各领域得分。上述协变量存在缺失数据的学生未纳入倾向得分估计。

权重计算如下：干预组权重 =  $1 / \text{倾向得分}$ ；对照组权重 =  $1 / (1 - \text{倾向得分})$ 。该方法构建了一个观测协变量分布独立于干预分配的加权样本，近似随机对照试验的条件。该方法使研究符合What Works Clearinghouse（WWC, 2024）“符合标准但有保留”的证据要求。WWC认可倾向得分匹配、分层和加权作为减少准实验设计偏差的有效策略。具体而言，倾向得分加权（如IPTW）可用于实现干预组与比较组的等值。

## 学业成就与Panorama合成得分的建模

我们采用双重差分（DiD）设计和固定效应线性回归模型，评估干预对学业成就和Panorama合成得分的影响。该方法通过纳入学生和学校层面的固定效应，控制未观测到且不随时间变化的个体特征，从而识别同一学生随时间的变化。主要自变量为干预指标、时间段指标及其交互项，用于评估干预组与对照组在指定时期内的差异性变化。模型纳入倾向得分权重以调整组间基线差异，并在学校层面聚类标准误，以考虑校内相关性。

我们使用以下双重差分模型估计干预效应：

$$Y_{it} = \beta_0 + \beta_1 \text{Treatment}_i + \beta_2 \text{Post}_t + \beta_3 (\text{Treatment}_i \times \text{Post}_t) + \alpha_i + \epsilon_{it}$$

其中， $Y_{it}$ 表示学生*i*在时间*t*的结果测量； $\text{Treatment}_i$ 是是否分配到干预组的二元指标； $\text{Post}_t$ 是干预后时期的二元指标； $\text{Treatment}_i \times \text{Post}_t$ 为交互项，体现干预效应的DiD估计； $\alpha_i$ 表示用于控制不随时间变化特征的个体固定效应； $\epsilon_{it}$ 为误差项。

表 2  
按时间列示合成变量的均值、标准差与内部一致性估计 (Cronbach' s  $\alpha$ )

| 时间      | 量表    | $\alpha$ | M    | SD   |
|---------|-------|----------|------|------|
| 2022年秋季 |       |          |      |      |
|         | 师生关系  | .868     | 3.60 | 0.95 |
|         | 学校归属感 | .875     | 3.83 | 0.99 |
|         | 学校氛围  | —        | 3.44 | 1.17 |
|         | 自我管理  | .842     | 4.03 | 0.62 |
|         | 支持性关系 | .634     | 1.84 | 0.21 |
| 2023年秋季 |       |          |      |      |
|         | 师生关系  | .856     | 3.70 | 0.91 |
|         | 学校归属感 | .875     | 3.93 | 0.96 |
|         | 学校氛围  | —        | 3.52 | 1.13 |
|         | 自我管理  | .840     | 4.06 | 0.62 |
|         | 支持性关系 | .635     | 1.86 | 0.20 |
| 2023年春季 |       |          |      |      |
|         | 师生关系  | .873     | 3.49 | 0.97 |
|         | 学校归属感 | .887     | 3.68 | 1.04 |
|         | 学校氛围  | —        | 3.26 | 1.19 |
|         | 自我管理  | .858     | 3.99 | 0.65 |
|         | 支持性关系 | .647     | 1.85 | 0.21 |
| 2024年春季 |       |          |      |      |
|         | 师生关系  | .878     | 3.62 | 0.95 |
|         | 学校归属感 | .891     | 3.92 | 1.01 |
|         | 学校氛围  | —        | 3.42 | 1.15 |
|         | 自我管理  | .869     | 4.04 | 0.66 |
|         | 支持性关系 | .687     | 1.86 | 0.21 |

表 3  
按时间与干预组别列示合成变量的均值和标准差

| 时间      | 组别  | TSR<br>M (SD) | SB<br>M (SD) | SC<br>M (SD) | SM<br>M (SD) | SR<br>M (SD) |
|---------|-----|---------------|--------------|--------------|--------------|--------------|
| 2022年秋季 | 对照组 | 3.59 (0.96)   | 3.80 (1.00)  | 3.45 (1.17)  | 4.07 (0.63)  | 1.84 (0.21)  |
|         | 干预  | 3.62 (0.94)   | 3.88 (0.97)  | 3.42 (1.16)  | 3.99 (0.61)  | 1.85 (0.21)  |
| 2023年春季 | 对照组 | 3.52 (0.96)   | 3.73 (1.03)  | 3.33 (1.18)  | 4.01 (0.65)  | 1.86 (0.20)  |
|         | 干预  | 3.44 (0.99)   | 3.57 (1.04)  | 3.12 (1.21)  | 3.97 (0.66)  | 1.84 (0.22)  |
| 2023年秋季 | 对照组 | 3.68 (0.92)   | 3.91 (0.98)  | 3.52 (1.14)  | 4.06 (0.62)  | 1.86 (0.20)  |
|         | 干预  | 3.75 (0.90)   | 3.97 (0.92)  | 3.54 (1.11)  | 4.07 (0.62)  | 1.86 (0.20)  |
| 2024年春季 | 对照组 | 3.60 (0.97)   | 3.91 (1.04)  | 3.46 (1.18)  | 4.03 (0.66)  | 1.86 (0.21)  |
|         | 干预  | 3.66 (0.91)   | 3.93 (0.94)  | 3.34 (1.08)  | 4.05 (0.65)  | 1.87 (0.21)  |

注. TSR = 师生关系；SB = 学校归属感；SC = 学校氛围；SM = 自我管理；SR = 支持性关系。

表 4  
Panorama学生调查的等值性检验

| 时间      | TSR   | SB    | SC    | SM    | SR    |
|---------|-------|-------|-------|-------|-------|
| 2022年秋季 | 0.03  | 0.08  | -0.03 | -0.13 | 0.05  |
| 2023年春季 | -0.08 | -0.15 | -0.18 | -0.06 | -0.10 |
| 2023年秋季 | 0.08  | 0.06  | 0.02  | 0.02  | 0.00  |
| 2024年春季 | 0.06  | 0.02  | -0.11 | 0.03  | 0.05  |

注. 所有统计量均为效应量d。TSR = 师生关系；SB = 学校归属感；SC = 学校氛围；SM = 自我管理；SR = 支持性关系。

我们还使用标准化均值差（SMD）检验协变量平衡，以评估加权策略是否成功平衡组间基线特征。SMD通过cobalt程序包的bal.tab()函数计算，并采用逆概率权重调整协变量差异。结果见表5。所有基线协变量，包括2022年秋季Panorama合成得分，均达到等值，即差异小于0.25个标准差单位。平衡分析纳入对照组1,585名学生和干预组750名学生。

该模型运用DiD方法比较干预组与对照组随时间的变化，有效控制可能混淆干预效应估计的未观测、时间不变因素。

对于Panorama合成得分模型，2022年秋季基线得分通过倾向得分权重纳入，因此符合WWC（2024）“符合标准但有保留”的基线等值要求。遗憾的是，我们没有学生个体层面的学业基线得分。为达到该标准，我们在学业模型中加入2021–2022学年（Second Step实施前一年）各校学生南卡罗来纳州大学与职业准备评估（SC READY）英语语言艺术和数学测试中达到或超过标准的比例。根据WWC，在缺少个体前测数据时，分析必须在以下方面达到基线等值：（a）宽泛、连续且标准化的学业成就测量；

(b) 至少两个人口学特征，如年级和种族/族裔（WWC, 2024, pp. 54–55）。表6列出各学科的学校层面均值和标准差。干预前组间差异为ELA的 $d = -0.06$ 和数学的 $d = -0.15$ ，均低于0.25的等值阈值。因此，纳入前一年学校层面学业表现满足基线要求。

### 行为与出勤结果的建模

我们估计一系列负二项回归模型，考察干预暴露和时间对ODR、ISS、OSS及到校天数等结果的影响。鉴于因变量为计数且存在过度离散，我们采用含分析权重和学校固定效应的负二项模型，即R语言MASS程序包的glm.nb函数。

每个模型均包括干预状态的二元指标（treat）、学年的分类变量（year）及二者交互项。模型还使用2021–2022学年学校层面ELA表现控制既往学业成就。固定效应设定控制未观测且时间不变的个体特征，识别同一学生随时间的变化。倾向得分权重用于调整干预组和对照组的基线差异，确保可比性。标准误在学校层面聚类，以考虑校内相关性，即同校学生可能呈现相关行为。

模型形式如下：

$$\log(\mu_i) = \beta_0 + \beta_1 \text{Treatment}_i + \beta_2 \text{Year}_i + \beta_3 (\text{Treatment}_i \times \text{Year}_i) + \beta_4 \text{ELA22}_i + \gamma_{s(i)} + \log(w_i)$$

其中， $\mu_i$ 为学生i的预期结果计数； $\text{Treatment}_i$ 是干预分配的二元指标； $\text{Year}_i$ 表示干预后学年； $\text{Treatment}_i \times \text{Year}_i$ 为体现DiD估计的交互项； $\text{ELA22}_i$ 为既往ELA得分； $\gamma_{s(i)}$ 为学校固定效应； $w_i$ 为分析权重（通过weights参数纳入）。

表 5  
协变量平衡

| 协变量          | 类型 | 标准化均值差 (SMD) |
|--------------|----|--------------|
| 六年级          | 二元 | 0.03         |
| 七年级          | 二元 | 0.02         |
| 八年级          | 二元 | -0.04        |
| 性别: 男性       | 二元 | -0.04        |
| 性别: 女性       | 二元 | 0.04         |
| 性别: 其他/未知    | 二元 | -0.00        |
| 种族: 白人       | 二元 | 0.21         |
| 种族: 黑人       | 二元 | -0.05        |
| 种族: 西班牙裔/拉丁裔 | 二元 | -0.16        |
| 种族: 亚裔       | 二元 | -0.01        |
| 种族: 美洲原住民    | 二元 | -0.00        |
| 种族: 太平洋岛民    | 二元 | 0.00         |
| 种族: 多种族      | 二元 | 0.01         |
| 英语学习者        | 二元 | 0.04         |
| 残障学生         | 二元 | 0.02         |
| 师生关系         | 连续 | 0.03         |
| 学校归属感        | 连续 | 0.07         |
| 学校氛围         | 连续 | -0.02        |
| 自我管理         | 连续 | -0.08        |
| 支持性关系        | 连续 | 0.05         |

注. 标准化均值差 (SMD) 绝对值超过0.10, 通常表示具有实际意义的不平衡。此处报告的数值均为加权后结果。

表 6  
2021–2022 学年学校层面的英语语言艺术与数学表现

| 组别  | ELA         | 数学          |
|-----|-------------|-------------|
|     | M (SD)      | M (SD)      |
| 干预  | 47.3 (11.9) | 37.1 (18.5) |
| 比较组 | 49.1 (40.0) | 42.1 (43.6) |

注. 数值为学生在南卡罗来纳州READY总结性评估中达到或超过标准的百分比。

为便于解释，我们对模型系数取指数，得到发生率比（IRR）。在负二项模型中，IRR表示其他变量不变时，预测变量每增加一个单位所对应的预期结果计数比例变化。例如，干预与学年交互项的IRR为1.20，表示该学年干预学校学生的结果计数（如到校天数或ODR）比较组高20%；IRR为0.80则表示预期计数减少20%。我们报告IRR及其95%置信区间，以反映估计的不确定性。随后，将IRR的自然对数除以对照组结果变量的估计标准差，转换为d（Sanchez-Meca et al., 2003）。我们分别对ODR、ISS、OSS和出勤建模，考察干预是否与行为和出勤随时间的改善相关。

结果

我们估计加权普通最小二乘（OLS）模型，预测2022–2023和2023–2024学年一年级ELA与数学成绩，以评估干预对学业结果的影响。各模型均控制相应学科的学校层面基线表现（2021–2022学年），纳入学校ID固定效应，并使用分析权重（wt）。

我们使用加权负二项回归，将学生行为结果（转介、ISS和OSS）作为计数变量建模（见表7）。每个模型均纳入干预与学年的交互项、基线ELA得分（ELA22）及学校固定效应，并仅分析所有纳入变量数据完整的个案。

表 7  
预测学生行为结果的负二项回归结果

| 结果变量 | 预测变量         | 估计值   | SE   | z     | p      |
|------|--------------|-------|------|-------|--------|
| 转介   | 截距           | 1.25  | 0.08 | 15.69 | < .001 |
|      | 干预           | -0.44 | 0.07 | -6.66 | < .001 |
|      | 2024年春季      | -0.12 | 0.07 | -1.70 | .089   |
|      | 干预 × 2024年春季 | -0.17 | 0.10 | -1.64 | .100   |
| ISS  | 截距           | -0.18 | 0.11 | -1.66 | .098   |
|      | 干预           | -0.37 | 0.10 | -3.73 | < .001 |
|      | 2024年春季      | -0.06 | 0.10 | -0.55 | .582   |
|      | 干预 × 2024年春季 | -0.45 | 0.15 | -3.00 | .003   |
| OSS  | 截距           | 0.65  | 0.10 | 6.58  | < .001 |
|      | 干预           | -0.22 | 0.11 | -2.10 | .036   |
|      | 2024年春季      | -0.19 | 0.11 | -1.73 | .084   |
|      | 干预 × 2024年春季 | -0.40 | 0.16 | -2.51 | .012   |

注. ODR = 纪律办公室转介；ISS = 校内停课处分；OSS = 校外停学处分。所有模型均调整基线英语语言艺术表现并纳入学校固定效应（因共线性省略参照类别）。负二项回归采用基于倾向得分的逆概率加权。

学业成就

如表8所示，即使调整学生基线学业表现后，干预（即忠实实施Second Step）仍与较高ELA得分显著相关。2022–2023学年，干预组平均比对照组高2.16分（SE = 0.51,  $p < .001$ ），对应 $d = 0.14$ 。2023–2024学年，这一优势略增至2.34分（SE = 0.49,  $p < .001$ ），相当于 $d = 0.17$ 。这些结果反映了干预组两年间幅度较小但一致的ELA提升。

表 8  
控制既往成就与学校固定效应后，预测英语语言艺术和数学成绩的加权普通最小二乘回归结果

| 结果变量      | 预测变量  | 估计值   | SE   | t      | p      | 调整后R <sup>2</sup> | RMSE | n     |
|-----------|-------|-------|------|--------|--------|-------------------|------|-------|
| 22-23 ELA | 截距    | 77.07 | 0.71 | 108.88 | < .001 | .142              | 15.8 | 2,951 |
|           | 干预    | 2.16  | 0.51 | 4.27   | < .001 |                   |      |       |
|           | 基线ELA | 0.14  | 0.01 | 12.25  | < .001 |                   |      |       |
| 23-24 ELA | 截距    | 76.54 | 0.73 | 104.29 | < .001 | .124              | 13.7 | 2,801 |
|           | 干预    | 2.34  | 0.49 | 4.78   | < .001 |                   |      |       |
|           | 基线ELA | 0.14  | 0.01 | 12.15  | < .001 |                   |      |       |
| 22-23 数学  | 截距    | 73.28 | 0.60 | 122.45 | < .001 | .110              | 15.6 | 2,937 |
|           | 干预    | 0.38  | 0.50 | 0.76   | .447   |                   |      |       |
|           | 基线数学  | 0.18  | 0.01 | 17.54  | < .001 |                   |      |       |
| 23-24 数学  | 截距    | 69.44 | 0.60 | 114.88 | < .001 | .185              | 13.1 | 2,795 |
|           | 干预    | 0.64  | 0.47 | 1.36   | .174   |                   |      |       |
|           | 基线数学  | 0.25  | 0.01 | 24.61  | < .001 |                   |      |       |

注. 所有模型均纳入学校固定效应（因共线性移除两个参照类别），并使用wt1变量进行分析加权。RMSE = 均方根误差；调整后R<sup>2</sup> = 调整后的决定系数。

相比之下，数学干预效应在两个学年均未达到统计学显著。2022-2023学年效应接近零（ $b = 0.38, p = .45; d = 0.02$ ）；2023-2024学年效应略大，但仍不显著（ $b = 0.64, p = .17; d = 0.05$ ）。这些效应量表明，两年间干预对数学成就几乎没有或没有可测量的影响。

行为结果

干预组总体ODR显著较少（ $b = -0.44, SE = 0.07, p < .001$ ）。干预与2024年春季的交互作用不显著（ $b = -0.17, p = .10$ ），提示转介效应在两个干预学年较为稳定。干预组总体ISS显著较少（ $b = -0.37, SE = 0.10, p < .001$ ），且2024年春季降幅更大（ $b = -0.45, p = .003$ ），提示实施两年后ISS减少效应增强。干预还与总体OSS较少相关（ $b = -0.22, SE = 0.11, p = .036$ ），且2024年春季交互项显著（ $b = -0.40, p = .012$ ），提示效应随时间增强。

为解释干预效应，我们将负二项回归系数取指数得到IRR，反映干预暴露所对应的预期行为事件计数的乘数变化。对于ODR，干预主效应表明，干预组转介次数比对照组约少36%（ $IRR = 0.64$ ，标准化均值差 $[d] = -0.15^1$ ）。与学年的交互作用虽不显著，但不同时点效应一致，提示转介减少模式较为稳健。对于ISS，2024年春季效应尤为明显；交互项表明，相比同期对照组，干预组ISS事件少36%（ $IRR = 0.64/d = -0.44^2$ ）。同样，OSS的干预与2024年春季交互作用显著，该学年干预组OSS事件少33%（ $IRR = 0.67, d = -0.42^3$ ）。这一模式提示，干预在第二年对较严重违规行为的减少作用有所增强。

<sup>1</sup> 比较组ODR标准差在2022–2023学年为3.23，2023–2024学年为2.85。  
<sup>2</sup> 比较组ISS标准差在2022–2023学年为1.01，2023–2024学年为1.03。  
<sup>3</sup> 比较组OSS标准差在2022–2023学年为0.99，2023–2024学年为0.90。

出勤

我们采用加权负二项回归预测到校天数，控制干预状态、学年、前一年ELA表现和学校固定效应（表9）。干预组总体到校天数显著高于对照组（ $b = 0.014$ ,  $SE = 0.002$ ,  $p < .001$ ）。这一主效应对应IRR = 1.014，即干预组平均到校天数多1.4%。按每学年180天计算，干预组比同伴约多到校2.5天。

干预与2024年春季交互作用不显著（ $b = 0.002$ ,  $p = .575$ ），提示积极出勤效应在两个实施学年保持稳定。2024年春季主效应也不显著（ $p = .666$ ）。基线ELA表现与出勤正相关（ $b = 0.00052$ ,  $p < .001$ ），说明学生学业表现较好的学校往往出勤更稳定。

表 9  
预测到校天数的负二项回归结果

| 预测变量          | 估计值   | SE    | z       | p      |
|---------------|-------|-------|---------|--------|
| 截距            | 5.09  | 0.003 | 1622.18 | < .001 |
| 干预            | 0.01  | 0.002 | 5.93    | < .001 |
| 2024年春季       | -0.00 | 0.003 | -0.43   | .666   |
| ELA22 (基线ELA) | 0.01  | 0.001 | 11.14   | < .001 |
| 干预 × 2024年春季  | 0.00  | 0.004 | 0.56    | .575   |

注. 模型纳入学校ID固定效应（因共线性剔除两个类别），并采用基于倾向得分的逆概率加权。ELA = 英语语言艺术。

Panorama各领域

我们估计DiD模型，考察Second Step对学校氛围感知和幸福感的影 响，重点分析2023年春季至2024年春季的变化。结果见表10。所有模型均包括学生和 学校固定效应，并采用倾向得分逆概率权重。干预与2024年春季交互项用于识别干预随时间产生的差异性影响。

表 10  
预测学校氛围与幸福感的双重差分 (DiD) 结果

| 结果变量       | 干预效应   | 2024年春季 | 干预 × 2024年春季 | RMSE | d     |
|------------|--------|---------|--------------|------|-------|
| 师生关系 (TSR) | 1.12*  | -0.08   | 0.22         | 0.54 | 0.40  |
| 学校归属感 (SB) | -0.66* | 0.01    | 0.21         | 0.55 | 0.37  |
| 学校氛围 (SC)  | -0.98* | 0.02    | 0.20         | 0.73 | 0.27  |
| 自我管理 (SM)  | 0.08   | 0.01    | 0.04         | 0.31 | 0.13  |
| 支持性关系      | 0.19   | 0.02    | -0.01        | 0.20 | -0.06 |

\* $p < .01$

注. “干预 × 2024年春季” 表示DiD估计, 即2024年春季干预组相对于对照组的额外变化。RMSE = 均方根误差; d = 交互项 ÷ RMSE。

若干学校氛围构念呈现统计学显著的正向DiD效应。相较对照组, 干预组师生关系 ( $b = 0.22, p = .001$ )、学校归属感 ( $b = 0.21, p = .003$ ) 和学校氛围 ( $b = 0.20, p = .031$ ) 均显著提升, 对应d分别为0.40、0.37和0.27, 提示学生对学校环境的感受随时间获得中等程度且具有实际意义的改善。自我管理 ( $b = 0.04, p = .26; d = 0.12$ ) 和支持性关系 ( $b = -0.01, p = .60; d = -0.06$ ) 未见显著DiD效应, 提示干预对个体内部调节或同伴支持网络的影响较小。

讨论

本研究全面评估Second Step Middle School Program对学业成就、行为事件、出勤及社会情感幸福感的影响。研究结果丰富了中学普遍性SEL项目的证据基础, 并凸显关注青春期早期学生学校经历多个维度的重要性。

与既有研究一致 (Durlak et al., 2011; Corcoran et al., 2018), 我们的结果表明, 忠实实施Second Step与学生发展成果的显著改善相关。尤其是, 干预学校学生连续两个学年ELA表现均有幅度较小但统计学显著的提升。数学效应虽不显著, 但ELA结果与SEL可能更紧密关联读写领域的文献一致 (Taylor et al., 2017)。这些发现支持如下理论: 情绪调节、自我效能和人际沟通等SEL能力, 有助于学生投入学业内容, 尤其是在读写要求较高的学科中。

除学业提升外, 干预还对行为结果产生稳健影响。Second Step学校学生的ODR、ISS和OSS均显著较少, 且效应在实施第二年增强。这些发现体现了持续接触SEL的累积收益, 并提示该项目可通过减少违规行为促进更积极的学校氛围。初中阶段行为风险较高 (Akos, 2006; Onetti et al., 2019), 因此这些结果尤其具有说服力。项目不仅减少较轻微的转介, 也减少OSS等更严重事件, 进一步显示其作为第一层级行为支持的潜力。

干预组出勤改善, 相当于每学年增加约2.5天。这丰富了目前数量有限但不断增长的、将SEL干预与学校投入改善联系起来的文献 (Jones et al., 2021)。慢性缺勤持续受到全国关注, 尤其是在COVID-19疫情之后, Second Step等SEL项目可能成为提高教学可及性和学生联结感的可行路径。

学生自报结果显示，师生关系感知、学校归属感和整体学校氛围显著改善。效应程度中等，提示Second Step学校学生随时间体验到更加肯定和支持性的教育环境。自我管理和支持性关系未见效应，可能反映天花板效应或实施时间相对较短。不过，关系性构念的显著提升回应了SEL的核心目标之一：营造温暖、包容且相互尊重的学校共同体。

这些发现带来若干启示。第一，忠实实施的普遍性SEL不仅有利于情感发展，也能改善具体学业和行为结果。第二，持续实施似乎尤为有效，第二年出现更强效应，提示应重视长期投入和持续项目，而非短期或一次性活动。第三，项目改善了ELA，却未影响数学。未来研究应探索不同学业领域是否对SEL呈现不同反应，以及学科特定调整能否增强学业效应。

### 局限与未来方向

尽管本研究具有大样本、采用倾向得分加权的严谨准实验设计及整合多来源数据等优势，仍存在局限。第一，加权提高了组间可比性，但不能完全排除未测量混淆变量。未来应采用随机对照试验增强内部效度。第二，缺少个体学业基线数据，使学业模型的等值检验不得不依赖学校层面成就替代指标。第三，样本的地理和人口学特征可能不同于其他实施Second Step的学区，限制结果的可推广性。

未来研究应探索SEL影响学业和行为结果的中介机制，如学校氛围、动机或执行功能变化。还应考察特定学生群体是否获益更多，尤其是残障学生、多语学习者和长期缺勤学生。鉴于政策日益重视全人教育，还需进一步评估中学SEL项目的成本效益、可扩展性和可持续性。

### 结论

本研究结果支持Second Step Middle School Program在多层级支持系统内促进初中生学业投入、改善行为结果及营造更积极学校经历的有效性。尽管初中常伴随风险上升和动机下降，本研究进一步丰富了相关文献，表明普遍性SEL可作为保护因素，缓冲发展风险，并为学习、归属和成长创造条件。

## 作者说明

Lauren Evanovich <https://orcid.org/0000-0001-9518-0259>

Nicholas A. Gage <https://orcid.org/0000-0002-1800-1760>

Robin Parks Ennis <https://orcid.org/0000-0002-8620-4199>

Tracy Ulrich <https://orcid.org/0009-0006-2354-2904>

作者无需要披露的利益冲突。

## 参考文献

- Akos, P. (2006). Middle school transition: Navigating change, keeping connected. *Middle School Journal*, 38(1), 5–12. <https://doi.org/10.1080/00940771.2006.11461527>
- Corcoran, R. P., Cheung, A. C. K., Kim, E., & Xie, C. (2018). Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research. *Educational Research Review*, 25, 56–72. <https://doi.org/10.1016/j.edurev.2017.12.001>
- Durlak, J. A., Mahoney, J. L., & Boyle, A. E. (2022). What we know, and what we need to find out about universal, school-based social and emotional learning programs for children and adolescents: A review of meta-analyses and directions for future research. *Psychological Bulletin*, 148(11-12), 765–782. <https://doi.org/10.1037/bul0000383>
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>
- Erentaitė, R., Vosylis, R., Gabrielavičiūtė, I., & Raižienė, S. (2018). How does school experience relate to adolescent identity formation over time? Cross-lagged associations between school engagement, school burnout and identity processing styles. *Journal of Youth and Adolescence*, 47(4), 760–774.
- Hornstra, L., Stroet, K., van Eijden, E., Goudsblom, J., & Roskamp, C. (2018). Teacher expectation effects on need-supportive teaching, student motivation, and engagement: A self-determination perspective. *Educational Research and Evaluation*, 24(3-5), 324–345. <https://doi.org/10.1080/13803611.2018.1550841>
- Jones, S. M., Bailey, R., Brush, K., & Kahn, J. (2021). Navigating SEL from the inside out: Looking inside and across 33 leading SEL programs. Harvard Graduate School of Education, EASEL Lab. <https://www.wallacefoundation.org/knowledge-center/pages/navigating-social-and-emotional-learning-from-the-inside-out.aspx>
- Niehaus, K., Rudasill, K. M., & Rakes, C. R. (2012). A longitudinal study of school connectedness and academic outcomes across sixth grade. *Journal of School Psychology*, 50(4), 443–460.

- Onetti, W., Fernández-García, J. C., & Castillo-Rodríguez, A. (2019). Transition to middle school: Self-concept changes. *PloS One*, 14(2), <https://doi.org/10.1371/journal.pone.0212640>
- Panorama Education. (n.d.). *Panorama social-emotional learning survey*. Available online at: <https://www.panoramaed.com/>
- Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, 8(4), 448–467. <https://doi.org/10.1037/1082-989X.8.4.448>
- Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, 88(4), 1156–1171. <https://doi.org/10.1111/cdev.12864>
- What Works Clearinghouse. (2024). *WWC standards handbook version 5.0*. U.S. Department of Education, Institute of Education Sciences. [https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc\\_standards\\_handbook\\_v5.pdf](https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v5.pdf)