

# 参与学生社团对大学生批判性思维能力的影响

——基于全国 78 所高校的实证研究

李文平

(郑州大学 教育学院,河南 郑州 450001)

**摘要:**学生社团活动是促进大学生批判性思维能力等综合素质提升的重要途径。基于全国 16 省 78 所高校的 4658 名大四本科生的能力测评数据,使用倾向得分匹配方法控制样本自选择偏误,定量评估参与学生社团给大学生带来的批判性思维能力收益。研究发现:社团参与对大学生批判性思维能力具有正向影响,且未参与社团学生潜在的批判性思维能力提升显著高于参与社团学生的提升;以领导团队成员、骨干成员身份参与社团活动的学生获得了显著的批判性思维能力提升,而仅作为一般参与者的学生获得的提升并不显著。

**关键词:**大学生社团;批判性思维能力;倾向得分匹配

**中图分类号:**G645.5 **文献标识码:**A **文章编号:**2095-9303(2023)01-0099-10

## 一、问题提出与文献回顾

作为一种获取知识和决定合理行动的思维方式<sup>[1]5</sup>,批判性思维与人的开放性、超越性联系在一起,是人类综合素质的重要表现<sup>[2]</sup>。许多证据表明它与创造力、问题解决能力等高阶能力紧密相关<sup>[3-7]</sup>,美国、英国、埃及、菲律宾等国家都将培养批判性思维能力作为高等教育的重要目标<sup>[8]</sup>。作为实施素质教育的重要途径和有效方式,高校学生社团活动在加强校园文化建设、提高学生综合素质、引导学生适应社会、促进学生

成才就业等方面发挥着重要作用<sup>[9]</sup>。诸多研究认为,社团活动对培养大学生批判性思维能力、领导能力及实现个性化发展等具有积极意义<sup>[10-12]</sup>。进而,如何利用学生社团活动培养大学生批判性思维能力受到了研究者的关注,但所得结论并未达成一致。

一些研究显示,学生社团活动显著提升了大学生批判性思维能力。一项针对 9 个国家大学生参与辩论活动的实证研究显示,参与辩论活动的学生,其批判性思维能力显著高于未参与的学生<sup>[13]</sup>。刘航等使用加利福尼亚批判性思维技能测试(CCTST)获得的前测与后测结果显示,参加

收稿日期:2022-06-02

基金项目:河南省教育科学“十四五”规划 2022 年度重点课题“基于第二课堂的大学生批判性思维能力培养研究”(2022JKZD31)

作者简介:李文平(1988—),男,郑州大学教育学院讲师,博士,研究方向:学生学习与发展、教育政策监测与评价。

辩论活动显著提高了学生的评价、推论、归纳推理和演绎推理技能<sup>[14]</sup>。阿斯汀等人通过研究大学生参与社区服务对自身发展产生的影响,发现参与过社区服务的学生在批判性思维能力、写作能力和成绩方面取得了显著的提升<sup>[15]</sup>。马慧等人采用准实验设计的方法发现,参与志愿服务活动对大学生的批判性思维技能具有显著促进作用,并且这种促进作用可以迁移到其他领域的批判性思维表现中<sup>[16]</sup>。帕斯卡雷拉认为,参与社团活动之所以会对批判性思维能力产生积极影响,部分原因在于社团活动增加了学生与不同背景的同龄人交往的可能性,增加了他们接触不同观点、不同价值观的机会<sup>[17]</sup>。

另有研究发现,参与社团活动并未显著提升大学生批判性思维能力,甚至产生了抑制作用。帕斯卡雷拉等人使用美国大学生学术水平评估(CAAP)中的批判性思维模块,探讨了参与兄弟会或姐妹会对大学生批判性思维能力的影 响。在控制混杂变量影响后发现,大一学年结束时,在批判性思维能力分数表现层面,兄弟会成员显著低于非成员,姐妹会成员虽然分数低于非成员,但差异不显著<sup>[18]</sup>。马丁等人使用前测、后测设计,并控制各种混杂因素影响后发现,兄弟会成员与非成员在各种教育结果(批判性思维能力、GPA等)的表现上未产生显著差异<sup>[19]</sup>。哈耶克认为,兄弟会成员需要对组织做出更多的时间和精力承诺,这会使他们难以平衡自己的组织责任和学习活动,进而不利于思维能力发展<sup>[20]</sup>。

无法有效识别社团参与影响批判性思维能力发展的因果效应是已有研究结论互斥的重要原因之一。已有研究多采用分组比较或者区组回归的方式,估算“参与社团”与“未参与社团”学生在批判性思维能力上的均值差异或(偏)回归系数差异。这类估计无法反映出因果效应,原

因有二:一是估计社团参与对批判性思维能力影响的数据是非完全随机抽样的,即无法保证参与社团的学生是随机分布的,基于非完全随机抽样数据的分组估计存在样本选择偏差<sup>[21]</sup>;二是存在混淆因素的干扰,即无法观察到批判性思维能力的提升是社团参与的影响还是研究对象自身特质等因素的贡献<sup>[22]</sup>。即便使用偏回归系数将“混杂因素”控制起来,仍存在三种估计偏误:第一,它假定混杂因素对批判性思维能力的影 响与社团参与对批判性思维能力的影 响之间是线性关系,这种假定缺乏科学依据<sup>[23]</sup>;第二,偏回归系数表示的是“平均”效应,它混淆了两个问题,即一个任意选取的参加社团的学生,若未参加社团会有怎样的批判性思维能力,以及一个任意选取的未参加社团的学生,如果参加了学生社团将会有怎么样的批判性思维能力;第三,混杂因素可能与我们关心的自变量“社团参与”存在高度相关。一般来说,完全随机试验是进行因果推断的理想方法<sup>[24]</sup>,但由于完全随机试验面临的反事实的不可观测性等问题<sup>①</sup>,倾向得分匹配(PSM)受到了研究者的关注和应用<sup>[25]</sup>。倾向得分匹配通过控制研究样本接受某种“处理”的概率来遏制样本自选择问题对因变量的影响,从而保证估计结果具有因果效应,同时它还通过一种半参数性方法规避了偏回归系数的三种估计偏误<sup>[26]</sup>。

综上所述,本研究使用倾向得分匹配方法控制样本自选择问题,基于全国16省78所院校的4658名大四本科生的能力测评数据,通过评估社团参与对大学生批判性思维能力影响的平均处理效应,回答研究命题。此外,本研究还希望了解学生社团参与程度对批判性思维能力影响的差异,以更深入地阐释研究命题,进而为相关研究提供资料积累,为本科生人才培养改革提供

①即任意一个研究样本*i*如果进入了试验组就不再可能同时进入对照组,故而样本*i*进入对照组这一“反事实”是无法观测的。

实证依据。

## 二、数据来源与变量设计

### (一) 数据来源

本文使用的数据源于“全国本科生能力测评”(NACC)。NACC由华中科技大学沈红教授带领的国家自然科学基金项目团队通过邀请国内外多学科领域专家研制、开发,包括能力测评和信息自评两部分<sup>①</sup>。其中,批判性思维能力模块由加籍华裔学者组织的批判性思维研究专家团队专门针对中国大学生开发、调试,模块由33道单项选择题构成,各题设计4—5个选项。NACC调查团队采用严谨的程序设计,于2016年12月以实地、发放纸质试卷、集中学生,并尽可能让参与院校按文、理、工、医学科类别各抽取50人的方式,对全国16省(自治区/直辖市)83所抽样院校的16000余名本科生进行了测评。经数据清洗,NACC最终获得有效样本15336个,可用于本研究倾向得分匹配的大四学生样本共计4658个。

### (二) 变量设计

1. 因变量。本研究的因变量为批判性思维能力。使用NACC的批判性思维模块,从分析段落的论证结构、识别隐含假设、澄清意义、评估信息可信性、评估从给定陈述到结论的推理、评估信息可推出含义等6个维度测评了批判性思维能力。按每题3.03分予以赋值,将33题折算为百分制统计。此变量作连续变量处理。

2. 干预变量。本研究使用的干预变量为大学生社团参与。根据共青团中央和教育部在《关于加强和改进大学生社团工作的意见》(中青联发[2005]5号)中的定义,“大学生社团是由高校学生依据兴趣爱好自愿组成,按照章程自主开展

活动的学生组织”。本研究涉及的学生社团共包括6种类别。鉴于本研究关注了两个问题,即社团参与能否显著提升大学生的批判性思维能力,以及社团参与程度对批判性思维能力影响的差异,本文使用NACC中的2个问题对社团参与进行取值:第一个问题是“大学期间你参与过哪些类型的社团”,选项为“未参与”“专业学术类”“艺术兴趣类”“体育健身类”“公益服务类”“职业准备类”“其他类”。该题为多项选择题,将“未参与”编码为0(对照组),剩余选项中任选一项即为样本有社团参与经历,编码为1(干预组),此为干预1。第二个问题是“你在社团中承担过的最高职务为”,选项为“领导团队成员”“骨干成员”“一般参与者”。以学生担任的角色来衡量社团参与程度的理由在于,通常来说,较为重要的组织角色往往被赋予较高的组织期望、责任等。个体扮演的角色越重要,角色要求个体承担的责任、应尽的义务也就越多,个体参与组织建设、组织活动的程度也会越深。为清晰地观察不同参与程度所产生的平均处理效应,本文将第二个问题分解为3种干预方式:领导团队成员=1(干预组),未参与社团=0(对照组),此为干预2;骨干成员=1,未参与社团=0,此为干预3;一般参与者=1,未参与社团=0,此为干预4。

3. 协变量。协变量主要选取研究样本不随是否参与社团而改变,同时又会对参与社团和批判性思维能力产生影响的个体特征、家庭背景、高中经历、院校特征等前定变量。主要包括四种类别:第一,个体特征变量,包括性别(女=1、男=0)、民族(汉族=1、少数民族=0)、第一代大学生(是=1、否=0)、独生子女(是=1、否=0);第二,家庭背景变量,包括读大学前家庭所在地(省会或直辖市=1、其他城市=0)、家庭经济状

<sup>①</sup>更多关于测评的介绍见沈红2017年发表于《高等教育研究》第11期的“我国大学生的能力水平与高等教育增值”一文。

况(5级连续编码)、父母受教育程度(教育年限连续编码)、父母职业(4层次连续编码)<sup>①</sup>;第三,高中就学经历,包括高中所在地(省会或直辖市=1、其他城市=0)、高中类别(重点高中=1、普通高中=0)、高中班级类别(重点班=1、普通班=0)、高三成绩排名指数(成绩排名/班级人

数);第四,院校特征,包括院校类型(985大学=1、211大学=2、普通四年制大学=3、普通四年制学院=4)、学科(文=1、理=2、工=3、医=4),均作名义变量处理。

对以上三类变量均给予描述性分析处理,结果详见表1。

表1 变量的描述性分析

| 类别   | 变量       | 样本量   | 均值     | 标准差    | 最小值 | 最大值 |
|------|----------|-------|--------|--------|-----|-----|
| 因变量  | 批判性思维能力  | 4658  | 55.182 | 13.028 | 0   | 100 |
| 干预变量 | 社团参与     | 4543  | 0.820  | 0.382  | 0   | 1   |
|      | 领导团队     | 4442  | 0.396  | 0.489  | 0   | 1   |
|      | 骨干成员     | 4442  | 0.168  | 0.374  | 0   | 1   |
|      | 一般参与者    | 4442  | 0.253  | 0.434  | 0   | 1   |
| 协变量  | 性别       | 4624  | 0.513  | 0.499  | 0   | 1   |
|      | 民族       | 4636  | 0.921  | 0.271  | 0   | 1   |
|      | 第一代大学生   | 4658  | 0.653  | 0.475  | 0   | 1   |
|      | 独生子女     | 4599  | 0.374  | 0.484  | 0   | 1   |
|      | 大学前家庭所在地 | 4559  | 0.104  | 0.305  | 0   | 1   |
|      | 家庭经济状况   | 4541  | 2.550  | 0.743  | 1   | 5   |
|      | 父亲受教育程度  | 4553  | 10.806 | 3.753  | 0   | 22  |
|      | 母亲受教育程度  | 4538  | 9.544  | 4.128  | 0   | 22  |
|      | 父亲职业     | 4516  | 3.427  | 0.583  | 1   | 4   |
|      | 母亲职业     | 4522  | 3.324  | 0.615  | 1   | 4   |
|      | 高中所在地    | 4586  | 0.128  | 0.334  | 0   | 1   |
|      | 高中类别     | 4508  | 0.621  | 0.485  | 0   | 1   |
|      | 高中班级类别   | 4496  | 0.497  | 0.500  | 0   | 1   |
|      | 高中排名指数   | 4161  | 0.291  | 0.285  | 0   | 1   |
|      | 院校类别     | 4658  | 2.748  | 0.989  | 1   | 4   |
| 所在学科 | 4658     | 2.129 | 1.001  | 1      | 4   |     |

### 三、估计方法与分析步骤

#### (一) 倾向得分匹配法的基本框架

为纠正干预效应估计时存在的自选择偏差,罗森鲍姆和鲁宾创建了倾向得分匹配方

法,已被广泛用于政策效果评估、教育成果评估等领域。倾向得分匹配方法是在反事实框架下,依据被研究样本接受某种“处理”的虚拟条件或概率对其进行替换的一种分析方法<sup>[27]</sup>。该方法的基本逻辑是,设置一个反事实组(对照组),使其样本尽可能与事实组(干

<sup>①</sup>本研究借鉴陆艺学的社会阶层划分,以及沈红教授对职业层次的划分,将NACC设计的11个职业从高至低分为4个层次:国家和社会管理者、经理级人员和私营业主为最高层次(编码为4),专业技术人员(教师、医生、军人)和个体工商户为次高层次(编码为3),企事业单位普通员工和商业服务人员为第三层次(编码为2),产业工人、农民工、农业劳动者、城乡无业/失业/半失业者为第四层次(编码为1)。

预组)样本在可观测协变量上分布趋同。倾向得分指被研究个体在给定可观测协变量条件下受到某种“处理”的概率。依据此得分,将干预组个体与得分相同或相似的对照组个体进行匹配,以达到类似随机试验的设计。根据匹配后两组样本在结果变量上的平均差异,获得被研究个体在接受某种干预后的平均处理效应。

假定社团参与为干预变量  $D_i$ ,将本研究中的样本分为两组:参与过社团的干预组( $D_i = 1$ )和未参与过社团的对照组( $D_i = 0$ )。对于学生  $i$ ,其潜在结果变量批判性思维能力被定义为  $Y_i(D_i)$ ,则参与过社团的学生其平均处理效应(ATT)为:

$$\widehat{ATT} = \frac{1}{N_1} \sum_{i: D_i=1} (y_{1i} - \hat{y}_{0i}) \quad (1)$$

式(1)中  $N_1$  为干预组样本数,  $\sum_{i: D_i=1}$  表示仅对干预组个体进行加总,  $y_{1i}$  表示干预组学生可观测到的批判性思维能力,  $\hat{y}_{0i}$  表示干预组学生若未参加学生社团时的潜在批判性思维能力,这即为现实中无法观测到的“反事实批判性思维能力”。

同理,可为对照组的每个学生干预组中寻找合适匹配,未参与过社团的学生其平均处理效应(ATU)为:

$$\widehat{ATU} = \frac{1}{N_0} \sum_{i: D_i=0} (\hat{y}_{1i} - y_{0i}) \quad (2)$$

式(2)中  $N_0$  为对照组样本数,  $\sum_{i: D_i=0}$  表示仅对对照组个体进行加总,  $y_{0i}$  表示对照组学生可观测到的批判性思维能力,  $\hat{y}_{1i}$  表示对照组学生如果参加学生社团其潜在的批判性思维能力,该数值亦为无法观测到的“反事实批判性思维能力”。全体样本的平均处理效应(ATE)为:

$$\widehat{ATE} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_{1i} - \hat{0}_{0i}) \quad (3)$$

式(3)中  $N$  为样本容量;如果  $D_i = 1$ ,则  $\hat{y}_{1i} = y_{1i}$ ;如果  $D_i = 0$ ,则  $\hat{y}_{0i} = y_{0i}$ 。

## (二)倾向得分匹配法的步骤

倾向得分匹配方法主要分为三步<sup>[28]</sup>:一是预测倾向值。这一步是利用可观测协变量(组)“X”,使用 probit 模型或 logistic 模型来预测个体参与学生社团的概率。二是利用倾向得分进行匹配。常见的匹配方法有寻找倾向值最邻近的  $K$  个不同组个体的“ $K$  临近匹配”,限制倾向值得分绝对距离的“半径匹配”,根据核函数计算匹配权重的“核匹配”,以及使用局部线性回归来估计匹配权重的“局部线性回归匹配”。各项匹配方法的详细规则相关文献已有详细解释<sup>[29]</sup>,这里不再展开。三是基于匹配样本进行因果系数估计。利用匹配后样本,通过观测平均处理效应便可估计学生社团参与对大学生批判性思维能力的影

## 四、计量结果

### (一)传统 OLS 回归估计结果

在进行匹配之前,以简单分组方式对研究样本进行的均值显著性差异的检验结果如表 2 所示。就干预 1 来看,参与过学生社团的大学生批判性思维能力均值为 55.923 分,显著高于对照组得分(51.799 分);就干预 2—4 来看,是领导团队成员或骨干成员的学生,批判性思维能力得分显著高于对照组得分,但作为一般参与者的学生,批判性思维能力得分显著低于对照组得分。这说明,社团参与并不总是正向作用于大学生批判性思维能力发展。然而,这种简单比较是在未控制混杂因素影响下得出的,结果可能存在偏差。

表2 社团参与差异与批判性思维能力得分差异 T 检验

| 干预手段   |             | 倾向值得分匹配前 |          | 差分       | T 检验   |
|--------|-------------|----------|----------|----------|--------|
|        |             | 干预组(参与)  | 对照组(未参与) |          |        |
| 学生社团   | 是否参与(干预1)   | 55.923   | 51.799   | 4.124*** | 7.906  |
|        | 领导团队成员(干预2) | 56.943   | 54.148   | 2.795*** | 7.077  |
| 担任职务类别 | 骨干成员(干预3)   | 56.028   | 55.101   | 0.927*   | 1.785  |
|        | 一般参与者(干预4)  | 54.585   | 55.485   | -0.900** | -2.074 |

注: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ 。下同。

表3 社团参与对批判性思维能力影响的 OLS 回归

|                   | 模型1      | 模型2      | 模型3    | 模型4    |
|-------------------|----------|----------|--------|--------|
| D 学生社团经历(干预1)     | 0.268*** | —        | —      | —      |
| D 领导团队成员(干预2)     | —        | 1.279*** | —      | —      |
| D 骨干成员(干预3)       | —        | —        | 0.833* | —      |
| D 一般参与者(干预4)      | —        | —        | —      | 0.059  |
| 协变量               | Y        | Y        | Y      | Y      |
| R-Square          | 0.2093   | 0.2032   | 0.2015 | 0.2009 |
| Adjusted R-Square | 0.2041   | 0.1977   | 0.1960 | 0.1954 |
| F                 | 39.83    | 36.73    | 36.34  | 36.21  |

注:(1)D开头的变量为虚拟变量;(2)为准确考察各干预手段(虚拟变量)对批判性思维能力带来的影响,在考察某一干预手段的影响时,并未将其他干预手段作为控制变量,以“—”表示该虚拟变量未纳入 OLS 模型;(3)囿于篇幅,表中未详细呈现各控制变量的拟合结果,仅以“Y”表示 OLS 模型加入了控制变量;(4)协变量包括:个体特征、家庭背景、高中经历和院校特征。

由表3可知,在控制混杂因素(协变量)影响后,有社团经历的学生比没有社团经历的学生批判性思维能力得分显著高出0.268分;而就不同干预手段来看(干预2—4),是领导团队成员、社团骨干成员的学生,批判性思维能力得分分别比对照组学生得分显著高出1.279分和0.833分,仅作为一般参与者的学生批判性思维能力得分高出对照组学生0.059分,但并不显著。换言之,无论大学生在社团中如何任职,他们的批判性思维能力得分均高于对照组得分,且参与程度更深的学生得分显著更高。这说明,学生社团参与对大学生批判性思维能

力发展具有积极意义,更深度的社团参与对大学生批判性思维能力的贡献更大。值得注意的是,即便使用 OLS 回归仍无法避免上文提出的三种估计偏误,难以保证所得结论具有因果效应,因此需借助 PSM 进一步控制自选择偏误。

## (二) 倾向得分匹配效果检验

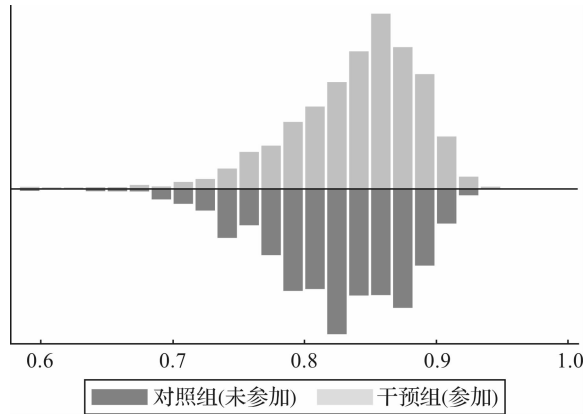


图1 干预组与对照组倾向值得分范围分布

为保障倾向得分匹配的估计效度,研究者需在估计平均处理效应之前,需对倾向得分匹配的效果进行共同支撑检验和条件独立假设检验。图1反映的是干预组和对照组学生倾向得分的共同取值范围。从图1可知,两组样本的倾向得分在[0.6,1]之间都有共同取值,并且多数样本的倾向得分落在[0.7,0.9]之间,表明干预组和对照组的倾向得分在共同取值域内有较大的重叠范围,能够满足共同支撑假设。此外,由图2可知,进行倾向得分匹配后,干预组和对照组样本在各协变量上的标准化偏差均在10%以内,表明

匹配后干预组和对照组样本在协变量上的差异得到了有效控制,满足了条件独立假设,能保证倾向得分匹配的有效性。

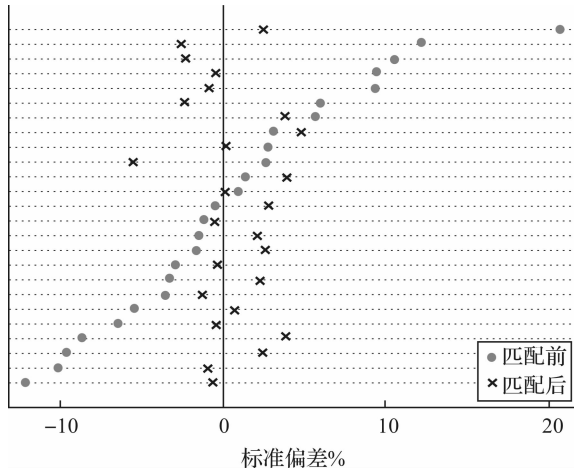


图2 倾向得分匹配前后协变量标准偏差对比

### (三) 社团参与对批判性思维能力影响的处理效应

表4 呈现的是不同匹配方法下社团参与(干预1)对大学生批判性思维能力影响的处理效应。以K临近匹配方法(K=1)为例,ATT的估计系数为1.017,表示从干预组随机选择一名学生,若他一开始就未参加社团,参加社团会使其批判性思维能力提升1.017分,但通过500次bootstrap抽样估计后发现,这种提升未呈现统计显著意义;ATU的估计系数为2.785,表示从对照组随机选择一名学生,若他参与了学生社团,其批判性思维能力将会提升2.785分,且500次bootstrap抽样估计显示此数值具有统计显著性;ATE的估计系数为1.305,表示从研究样本中随机选择一名学生,如果他参与了学生社团,其批判性思维能力会提升1.305分,且提升显著。综合来看,4种匹配方法下获得的平均处理效应估计值均为正值,并且所有的ATU估计值和ATE估计值均统计显著,所有的ATT估计值均未呈现统计显著意义;从估计系数的大小看,4种匹配方法的估计系数均呈现ATU > ATE > ATT的

特点。这意味着,其一,社团参与对大学生批判性思维能力具有正向影响,尤其是对于对照组学生而言,社团参与能显著提升其批判性思维能力;其二,社团参与给对照组学生带来的潜在批判性思维能力提升大于对干预组学生的提升。

表4 社团参与对大学生批判性思维能力影响的处理效应

| 匹配方法            | 处理效应 | 系数       | 标准误   | Z值   | P >  Z |
|-----------------|------|----------|-------|------|--------|
| 临近匹配<br>(k = 1) | ATT  | 1.017    | 0.783 | 1.17 | 0.241  |
|                 | ATU  | 2.785*** | 0.981 | 2.59 | 0.009  |
|                 | ATE  | 1.305*   | 0.708 | 1.67 | 0.095  |
| 半径匹配<br>(0.01)  | ATT  | 1.388    | 0.759 | 1.21 | 0.227  |
|                 | ATU  | 1.671**  | 1.043 | 2.44 | 0.015  |
|                 | ATE  | 1.434*   | 0.701 | 1.69 | 0.091  |
| 核匹配<br>(0.06)   | ATT  | 1.421    | 0.783 | 1.17 | 0.241  |
|                 | ATU  | 1.550**  | 1.028 | 2.47 | 0.013  |
|                 | ATE  | 1.442*   | 0.706 | 1.68 | 0.094  |
| 局部线性回归匹配(0.8)   | ATT  | 1.308    | 0.796 | 1.15 | 0.249  |
|                 | ATU  | 1.601*** | 0.958 | 2.66 | 0.008  |
|                 | ATE  | 1.356*   | 0.704 | 1.68 | 0.093  |

注:标准误、Z值和显著性通过bootstrap500次抽样后获得。下同。

### (四) 社团参与程度对批判性思维能力影响的处理效应

为验证本研究关注的第二个问题,我们以大学生在学生社团中担任的角色作为社团参与程度的操作变量,并将其分解为3种干预方式,即领导团队成员、骨干成员和一般参与者。同样使用4种匹配方法,评估不同干预手段下参与社团(干预组)与不参与社团(对照组)大学生的批判性思维能力差异状况。估计结果如表5所示。

在不同匹配方法下,以担任不同角色划分的干预变量的处理效应呈现了稳健的结论:第一,担任同种角色,使用不同匹配方法的估计结果显示,是社团领导团队成员或骨干成员的样本在4种匹配方法下获得的ATT、ATU、ATE估计系数均为正值,并且500次bootstrap抽样获得的ATU、ATE系数均显著,ATT系数均不显著;担任一般参与者的样本,虽然所有匹配方法下的

ATE、ATU 估计系数均为正值,但 ATT 估计值正负不一,并且用 4 种匹配方法获得的 3 种估计系数(ATT、ATU、ATE)均不显著。第二,同一匹配方法下,以担任不同角色为干预变量获得的处理效应也具有一致表现。使用半径匹配、核心匹配、局部线性回归匹配获得的 ATT、ATU、ATE 估计系数,均呈现出高度一致性特征,即领导团队成员 > 骨干成员 > 一般参与者。并且,3 种匹配方法得出的干预 2(领导团队成员)、干预 3(骨

干成员)估计系数(ATU、ATE)均呈现出统计学显著意义。此两点结论说明:其一,对于任意抽取的样本,参与学生社团都能正向作用于其批判性思维能力,这进一步验证了上文结论。其二,更深程度地参与学生社团才能显著作用于批判性思维能力提升,仅作为一般参与者参与学生社团对大学生批判性思维能力提升作用有限,并且可能对干预组样本(ATT)产生负向作用,阻碍其批判性思维能力发展。

表 5 社团参与程度对批判性思维能力影响的处理效应

| 干预手段            | 匹配方法          | 处理效应 | 系数        | 标准误   | Z 值  | P >  Z |
|-----------------|---------------|------|-----------|-------|------|--------|
| 领导团队成员<br>(干预2) | 临近匹配(k=1)     | ATT  | 0.628     | 0.632 | 1.57 | 0.116  |
|                 |               | ATU  | 1.386 **  | 0.615 | 2.37 | 0.018  |
|                 |               | ATE  | 1.079 *** | 0.487 | 2.60 | 0.009  |
|                 | 半径匹配(0.01)    | ATT  | 1.035     | 0.671 | 1.48 | 0.139  |
|                 |               | ATU  | 1.527 **  | 0.627 | 2.32 | 0.020  |
|                 |               | ATE  | 1.328 **  | 0.509 | 2.49 | 0.013  |
|                 | 核匹配(0.06)     | ATT  | 1.144     | 0.641 | 1.55 | 0.122  |
|                 |               | ATU  | 1.547 **  | 0.604 | 2.41 | 0.016  |
|                 |               | ATE  | 1.384 **  | 0.492 | 2.58 | 0.010  |
|                 |               | ATT  | 1.112 *   | 0.613 | 1.62 | 0.100  |
|                 |               | ATU  | 1.397 **  | 0.591 | 2.46 | 0.014  |
|                 |               | ATE  | 1.282 *** | 0.472 | 2.69 | 0.007  |
| 骨干成员(干预3)       | 临近匹配(k=1)     | ATT  | 1.741     | 0.903 | 1.16 | 0.248  |
|                 |               | ATU  | 1.023 *   | 0.698 | 1.74 | 0.081  |
|                 |               | ATE  | 1.145 *   | 0.629 | 1.89 | 0.059  |
|                 | 半径匹配(0.01)    | ATT  | 0.850     | 0.897 | 1.16 | 0.244  |
|                 |               | ATU  | 0.931 *   | 0.731 | 1.67 | 0.096  |
|                 |               | ATE  | 0.916 *   | 0.658 | 1.80 | 0.071  |
|                 | 核匹配(0.06)     | ATT  | 0.783     | 0.905 | 1.15 | 0.248  |
|                 |               | ATU  | 0.805 *   | 0.733 | 1.66 | 0.097  |
|                 |               | ATE  | 0.801 *   | 0.672 | 1.77 | 0.076  |
|                 | 局部线性回归匹配(0.8) | ATT  | 0.835     | 0.867 | 1.20 | 0.229  |
|                 |               | ATU  | 0.923 *   | 0.731 | 1.67 | 0.095  |
|                 |               | ATE  | 0.908 *   | 0.655 | 1.81 | 0.070  |
| 一般参与者(干预4)      | 临近匹配(k=1)     | ATT  | 0.551     | 0.723 | 0.35 | 0.730  |
|                 |               | ATU  | 0.114     | 0.658 | 0.43 | 0.667  |
|                 |               | ATE  | 0.223     | 0.563 | 0.49 | 0.625  |
|                 | 半径匹配(0.01)    | ATT  | -0.146    | 0.757 | 0.33 | 0.741  |
|                 |               | ATU  | 0.307     | 0.580 | 0.49 | 0.626  |
|                 |               | ATE  | 0.193     | 0.509 | 0.54 | 0.589  |
|                 | 核匹配(0.06)     | ATT  | -0.188    | 0.767 | 0.33 | 0.745  |
|                 |               | ATU  | 0.137     | 0.639 | 0.44 | 0.658  |
|                 |               | ATE  | 0.056     | 0.557 | 0.49 | 0.621  |
|                 | 局部线性回归匹配(0.8) | ATT  | -0.429    | 0.708 | 0.35 | 0.724  |
|                 |               | ATU  | 0.289     | 0.615 | 0.46 | 0.645  |
|                 |               | ATE  | 0.206     | 0.534 | 0.51 | 0.607  |



## 五、结论与讨论

基于“全国本科生能力测评”数据,使用倾向得分匹配方法,本研究定量估计了两个问题,一是参与学生社团能否提升大学生批判性思维能力,二是学生社团参与程度对批判性思维能力影响的差异。主要结论为:

第一,与传统 OLS 估计结果相比,倾向得分匹配方法获得的估计结果相对更高。无论是干预 1,还是干预 2—4,不同匹配方法获得的 ATE 系数大多比 OLS 估计系数更大。例如,干预 1 的 OLS 估计系数为 0.268,使用 4 种匹配方法获得的 ATE 系数均值为 1.384。这说明,参与学生社团对批判性思维能力影响的 OLS 模型中遗漏的潜在变量存在自相关可能,即 OLS 模型存在内生性风险,将造成估计结果有偏。这表明,本研究使用倾向得分匹配对社团参与的影响进行因果效应推断是有意义的。

第二,社团参与对大学生批判性思维能力具有正向影响。倾向得分匹配估计系数显示,使用 4 种匹配方法获得的平均处理效应 ATT、ATU、ATE 均为正值,并且所有匹配方法获得的 ATE 和 ATU 估计系数都是显著的。这意味着,参与社团所带来的批判性思维能力收益是正向的,并且对于目前未参与社团的大学生,参与社团将给他们带来更为显著的批判性思维能力提升。

第三,更深程度地参与社团才会对大学生批判性思维能力产生显著的正向影响。4 种匹配方法下,不同程度地参与社团给大学生带来的批判性思维能力处理效应均值(ATE 均值)从大到小依次是:领导团队成员(1.268)、骨干成员(0.942)、一般参与者(0.169)。并且,4 种匹配方法获得的领导团队成员、骨干成员的 ATE 估计系数均具有统计显著意义,一般参与者的 ATE 估计系数均未呈现统计显著意义。这意味着,对于任意抽取的研究样本,以领导团队成员、骨干成员身份参与社团活动的大学生获得了批判性

思维能力的显著提升,而仅作为一般参与者的的大学生获得的批判性思维能力提升并不显著。

上述发现从不同角度论证了一个总体结论:参与学生社团会正向影响大学生批判性思维能力,且较为深度地参与社团能显著促进大学生批判性思维能力。这种结果与批判性思维的社会性学习特征有关。概括地说,批判性思维的中心问题是论证<sup>[1]34</sup>,而合理论证的形成需要与他人对话,开展社会性学习<sup>[30]</sup>。通常而言,社团活动是群体性活动,甚至许多社团活动本身就是论证性的,需要参与者与其他成员开展较为深入的会话。这些会话不仅为参与者提供了练习论证的机会,也能让参与者更为全面、理性地认识议题,从而有利于其批判性思维能力发展<sup>[31]</sup>。

本研究是一项基于截面数据的事实推断研究,虽然使用了倾向值匹配方法来控制由个体特征、家庭背景、高中就学经历、院校特征等可观测协变量造成的异质性残差,但由于方法本身的局限性,无法纠正诸如智力禀赋、人格倾向、文化氛围等不可观测因素带来的影响,致使研究结论可能高估社团参与的作用。未来可基于追踪数据,采用倾向值匹配与双重差分相结合的方法,同时控制由可观测变量和不可观测变量带来的偏差,以图更为精确地估计社团参与给大学生批判性思维能力带来的增量。

### 参考文献:

- [1]董毓. 批判性思维原理和方法:走向新的认知和实践[M]. 高等教育出版社, 2011.
- [2]欧阳康. 批判性思维的前提性反思[J]. 高等教育研究, 2012(11):71-75.
- [3]ALEXANDER P A, JUDY J E. The Interaction of Domain-Specific and Strategic Knowledge in Academic Performance [J]. Review of Educational Research, 1988(58):375-404.
- [4]MARGINSON S. The problem of 'transferable' skills [J]. Critical Studies in Education, 1994(1):4-28.

- [5]江静,杨百寅.善于质疑辨析就会有高创造力吗:中国情境下的领导—成员交换的弱化作用[J].南开管理评论,2014(2):117-128.
- [6]胡卫平,ADEY P,申继亮,等.中英青少年科学创造力发展的比较[J].心理学报,2004(6):718-731.
- [7]CREBERT G, BATES M, BELL B, et al. Developing generic skills at university, during work placement and in employment: graduates' perceptions [J]. Higher Education Research & Development, 2004(2):147-165.
- [8]武宏志,周建武.批判性思维:论逻辑视角[M].北京:中国人民大学出版社,2010:22-32.
- [9]共青团中央,教育部.关于加强和改进大学生社团工作的意见[EB/OL].[http://www.ccyll.org.cn/documents/zqjf/200703/t20070321\\_14553.htm](http://www.ccyll.org.cn/documents/zqjf/200703/t20070321_14553.htm).
- [10]吴薇.职业服务类学生社团在大学生就业指导中的作用[J].思想理论教育,2008(7):73-77.
- [11]卢德平.青少年领导力的萌芽和形成[J].中国青年研究,2008(5):10-15.
- [12]李相阳,苑婧娴.浅谈高校社团对大学生个性化发展的影响[J].思想教育研究,2015(1):78-81.
- [13]COLBERT K. Enhancing critical thinking ability through academic debate [J]. Contemporary Argumentation & Debate, 1995(16):52-72.
- [14]刘航,金利民.英语辩论与大学生批判性思维发展的实证研究[J].外语与外语教学,2012(5):27-31.
- [15]ASTIN A W, VOGELGESANG L J, IKEDA E K, et al. How Service Learning Affects Students [EB/OL].<http://www.csufresno.edu/academics/cesl/documents/facultysl/facultysl-Astin-Exec-Summary.pdf>.
- [16]马慧,姚梅林,仝丽娟.服务学习促进大学生批判性思维的干预研究[J].心理发展与教育,2013(5):515-524.
- [17]PASCARELLA E T, PALMER B, MOYE M, et al. Do diversity experiences influence the development of critical thinking? [J]. Journal of College Student Development, 2001(3):257-271.
- [18]PASCARELLA E T, FLOWERS L, WHITT E J. Cognitive Effects of Greek Affiliation in College: Additional Evidence [J]. Journal of Student Affairs Research & Practice, 2009(3):815-836.
- [19]MARTIN G L, HEVEL M S, ASEL A M, et al. New Evidence on the Effects of Fraternity and Sorority Affiliation During the First Year of College [J]. Journal of College Student Development, 2011(5):543-559.
- [20]HAYEK J C, CARINI R M, O'DAY P T, et al. Triumph or Tragedy: Comparing Student Engagement Levels of Members of Greek-Letter Organizations and Other Students [J]. Journal of College Student Development, 2002(5):643-663.
- [21]HECKMAN J, ICHIMURAL H, SMITH J, et al. Characterizing Selection Bias Using Experimental Data [J]. Econometrica, 1998(5):1017-1098.
- [22]ROSENBAUM P R. The central role of the propensity scores in observational studies for causal effects [J]. Biometrika, 1983(1):41-55.
- [23]MORGAN S L. Counterfactuals, Causal Effect Heterogeneity, and the Catholic School Effect on Learning [J]. Sociology of Education, 2001(4):341-374.
- [24]胡安宁.社会科学因果推断的理论基础[M].北京:社会科学文献出版社,2015:4.
- [25]王舒鸿,崔欣,姚守宇.统计相关还是真实因果?——基于“因果推断”的新兴研究范式[J].金融与经济,2018(8):23-32.
- [26]蒋建忠.匹配与社会科学因果推论[J].实证社会科学,2017(2):29-48.
- [27]胡安宁.倾向值匹配与因果推论:方法论述评[J].社会学研究,2012(1):221-242.
- [28]ROSENBAUM P R, RUBIN D B. Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score [J]. The American Statistician, 1985(1):33-38.
- [29]陈强.高级计量经济学及Stata应用[M].北京:高等教育出版社,2010:544.
- [30]斯蒂芬·布鲁克菲尔德.批判性思维教与学[M].北京:中国人民大学出版社,2017:51.
- [31]李文平,沈红.学习性投入能带来批判性思维能力增值吗?[J].复旦教育论坛,2020(4):64-71.

(责任编辑:马速)