

批判性思维真的可教吗?

——基于79篇实验或准实验研究的元分析

冷静 路晓旭

(华东师范大学 教育学部教育信息技术学系,上海 200062)

[摘要] 批判性思维已成为21世纪人才必备的重要能力之一。关于批判性思维是否真的可教在学界一直存在争议。本研究选取79篇国内外实验或准实验研究文献为研究对象,采用元分析方法探究批判性思维的可教性。结果表明,所有学科领域和教育阶段都有提升学生批判性思维(包括技能和倾向)的有效教学策略。本研究从学科、学段、实验周期、实验人数、技术介入、国家区域和测评方式等方面探究批判性思维教学效果的差异。值得注意的是:1)批判性思维教学效果在不同学科差异显著,其中数学学科的效果最好;2)高中阶段开展批判性思维教学的效果最佳;3)批判性思维教学的周期越长,提升效果越好;4)教学干预在不同实验人数上具有同等的影响效果;5)有无技术介入对学生批判性思维的提升不存在显著差异;6)东西方国家的批判性思维教学效果存在显著差异;7)相较于标准化测试,自编工具测得的批判性思维能力的提升更明显。本研究采用元分析方法定量地综合了近十年的批判性思维实证研究,期待它能对未来批判性思维的培养提供启示。

[关键词] 批判性思维;元分析;教学干预

[中图分类号] G442

[文献标识码] A

[文章编号] 1007-2179(2020)06-0110-09

一、问题提出

批判性思维作为21世纪人才必备能力,有助于人们在信息社会开展有逻辑、独立的思考,甄别信息并快速作出决定(Reinstein et al., 2008)。批判性思维能力作为核心素养的重要组成部分,是我国教育现代化建设的战略重点(褚宏启,2016)。批判性思维能力包括分析、综合和评估辨别信息的技能以及运用这些技能的倾向(National Center for Excel-

lence in Critical Thinking Instruction, 1991)。对于学生来说,他们只有具备批判性思维能力,才能更好地理解和掌握学科知识,以及在真实的情境中运用这些知识解决具体问题(冷静等,2018)。

尽管人们普遍认为培养学生的批判性思维能力十分必要,但对批判性思维的可教性以及如何培养批判性思维能力的讨论一直存在争议。一方面,研究者认为通过教学提高学生的批判性思维,效果甚微(Mcmillan, 1987;Jonassen, 1999)。另一方面,研

[收稿日期] 2020-02-21

[修回日期] 2020-08-10

[DOI编码] 10.13966/j.cnki.kfjyyj.2020.06.011

[基金项目] 2017年度中央高校基本科研业务费项目华东师范大学人文社会科学青年预研究项目“互联网+环境下的学习焦虑如何克服?——基于大学生语言学习策略分析”(项目编号:2017ECNU-YYJ047)。

[作者简介] 冷静,博士,华东师范大学教育学部教育技术学系副教授,研究方向:计算机支持的协作学习、批判性思维能力培养、教育数据挖掘(jleng@deit.ecnu.edu.cn);路晓旭,华东师范大学教育学部教育技术学系硕士研究生,研究方向:学习分析、教育游戏。

[引用信息] 冷静,路晓旭(2020). 批判性思维真的可教吗?——基于79篇实验或准实验研究的元分析[J]. 开放教育研究, (6):110-118.

究者认为批判性思维可以通过设计特定而系统的教学来培养(Pascarella et al., 2006)。当前,国内外已有许多学校将批判性思维能力的培养纳入教学过程,并通过实证研究检验了教学干预对学生批判性思维的影响。然而,这些研究的实验结果差异较大,主要有以下两种截然相反的结论:一类研究结果表明,教学干预能有效促进批判性思维的提升。例如,有研究者以九年级学生为对象,在编程课中采用基于同伴评阅的在线教学策略学习。实验结果证明,此策略能够提高学生的批判性思维能力(Wang et al., 2017),再如,蒂鲁内等(Tiruneh et al., 2015)采用前测后测准实验设计在科学与技术课上实施系统设计“问题”教学策略,结果显示,实验组的得分显著高于控制组;另一类研究结果则表明,教学干预是无效的。例如,内伯等(Naber et al., 2014)利用反思性写作策略提高学生的批判性思维能力。然而,准实验结果证明不能显著提高学生的批判性思维能力。阿克于兹等(Akyüz et al., 2009)在课程设计课中采用混合式教学策略开展实验,结果表明,前后学生的批判性思维能力不存在显著差异,混合式教学策略无效。那么,批判性思维真的可教吗?教学干预的有效性是否受学段、学科、实验周期、实验人数、技术介入与否、国家区域以及不同测评方式的影响?仅仅个别的实证研究对总体上理解教学策略是否有效是没有帮助的。因此,本研究采用元分析方法分析教学干预对学生批判性思维的影响,并观察不同调节变量对批判性思维提升效果的影响。

综上,本研究分析与梳理了近十年国内外教学干预对学生批判性思维影响效果的 79 项实验与准实验研究,审视与评价教学干预对学生批判性思维能力的实际作用,以期对相关研究者了解有效的教学策略,开展后续研究与实践提供借鉴和指引。

二、研究设计

(一)文献搜集

为确保有效性和可靠性,本研究检索的文献为 Web of Science 核心数据库的期刊论文,以及中国知网(CNKI)收录的中文社会科学引文索引(CSSCI)来源期刊论文。Web of Science 的主题检索词包括:1)“critical thinking”和“pretest”或“posttest”;2)“critical thinking”和“control group”或“quasi exper-

iment”;分类选择 education, educational research, 和 education scientific disciplines, 文献类别选择 article, 检索时间选择 2009 年至 2019 年,检索得到 736 篇文献。研究者依据以下关键词在中国知网进行主题检索:1)“批判性思维”和“准实验”;2)“批判性思维”和“实验”;3)“批判性思维”和“前后测”,检索时间为 2009 年至 2019 年,检索得到 502 篇文献。

(二)文献纳入标准

由于检索到的文献并不全部符合要求,研究者对文献又进行了筛选:1)研究方法为实验研究或准实验研究,排除综述性文章及理论性文章;2)研究主题为教学方法或策略对学生批判性思维的影响;3)实验对象仅为在校学生,排除对象为成人学习者的文献;4)关注的是教学干预对学习批判性思维能力的提升效果,故文章应报告测量工具及评价指标;5)研究结果有充足的数据信息,如平均值、方差、t 值及 p 值等,有可靠的统计数据结果,可以据此计算平均效应值,排除无法计算效应值的文献;6)排除重复文献,若同一篇文章在不同期刊,或以不同形式发表,只取其一。样本筛选完成后,共有 79 篇文献符合标准。由于有的研究存在多组对照组与实验组,有的衡量了批判性思维技能与倾向两个方面,最终统计得到的有效效应值为 103 个(周榕等, 2019)。

(三)文献特征值编码

为了方便后期统计分析,本文对原始文献的各项特征值进行了编码,统计文献作者、年份、地区、学段、学科、样本量、实验周期、教学干预、测评工具。由于批判性思维分倾向和技能两个维度,因此本研究在教学效果上分倾向和技能两方面展开。本研究将学段分小学、初中、高中和大学四个阶段;在学科编码上,将效应值在 2 个及以上的列为独立学科,效应值不足 2 个且无法归类为上述课程的被列为其他学科(如会计、商店管理与规划、新闻文本阅读等);教学干预在多学科课程实施的列为综合课程,教学干预类型分有无技术介入两类;在国家区域编码上,按照东西方国家进行编码;在测评工具编码上,将经典的量表编码为标准化测试,其他为自编测验。为保证编码的准确性,两位编码者分别对 20 篇文献(占总数的 25%)进行预编码,编码结果的一致性为 0.85,说明编码的结果是可靠的。

(四)元分析方法

元分析法是一种量化分析法,能够根据以往的研究结果计算效应值,定量分析整体和系统,揭示同类研究的分歧与共性,从而得出普适性和规律性的结论(Bowman, 2012)。元分析可以用系统的方法将不同研究结果和矛盾的研究结果进行定量合成得出结论(王翠如等,2018)。本研究采用综合性元分析软件进行数据处理,计算效应值。效应值是一种标准化度量标准,表示两组平均值(通常是对照组和实验组)之间的差异。科恩(Cohen,1988)提出的效应值计算方法为:

$$d = \frac{\bar{X}_e - \bar{X}_c}{SD_{\text{Pooled}}}$$

这个基本方程具备两种情况的效应值计算:一种用于报告实验组和对照组的预测试数据研究,另一种用于单组的预测试后设计。在其他情况下(如t检验、F检验、p水平),效应量指标根据赫奇斯等人(Hedges et al., 2009)提供的转换公式估计。

由于不同研究实验或准实验的设计不同以及部分研究的样本量较小,本研究采用赫奇斯效应值估计平均差,以消除样本大小和不同实验设计对估计值的影响,函数关系式如下(Borenstein et al., 2009):

$$g \cong \left(1 - \frac{3}{4N-9}\right)d$$

标准误差用以下公式计算:

$$SE_g = \sqrt{\frac{1}{n_e} + \frac{1}{n_c} + \frac{g^2}{2N} \left(1 - \frac{3}{4N-9}\right)}$$

三、研究结果与讨论

(一)发表偏倚检验

本研究基于赫奇斯效应值计算出103个效应值的漏斗图(见图1)。横轴表示赫奇斯效应值,纵轴表示效应值的标准误。漏斗图的中上部分是有效区域,中线表示平均效应值。从图1可以看出,绝大多数研究的效应值落在漏斗图的上部有效区域内,且相对均匀地分布在漏斗图的平均效应值两侧,这说明存在出版偏差的可能性较小。以埃格(Egger)回归法和贝格(Begg)的秩相关法对发表偏倚进行检验,结果为: $t = 3.946, p < 0.001$; Begg检验结果为: $t = 0.243, p < 0.001$,说明可能存在发表偏倚。

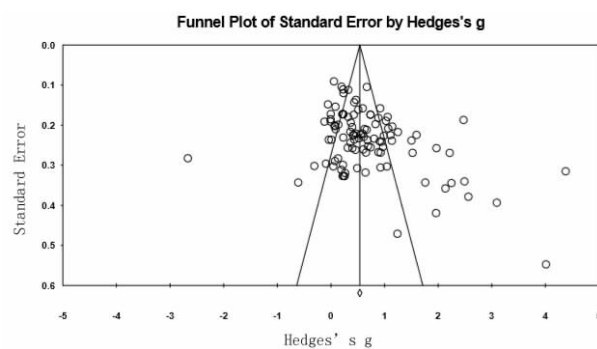


图1 元分析漏斗图

(二)异质性检验

异质性检验结果显示,Q值为1023.235,且统计结果达到显著水平($p < 0.001$),说明各项研究之间存在异质性。研究采用 I^2 统计量方法判断样本异质性程度。通常, $I^2 > 75\%$ 时采用随机效应模型, $0 \leq I^2 \leq 75\%$ 时采用固定效应模型消除异质性(Borenstein et al., 2009)。本研究中, $I^2 = 90\%$,即90%的异质性由效应值的真实差异造成,10%的异质性由误差导致。真实差异导致异质性出现的原因可能是:实验周期、研究学段、研究学科、国家区域、研究时间等。本研究中,由于79项元分析样本的异质性结果大于75%,故采用随机效应模型(Random Effects)评估教学干预对学生批判性思维提升的作用。

(三)教学干预对提升学生批判性思维整体效应的检验结果

1. 整体效应值

本研究通过计算79项研究的效应值及使用标准化测量作为结果变量,得出元分析教学干预对批判性思维的整体影响。科恩认为,当效应值在0.8以上时,教学干预对批判性思维提升的影响很大;当效应值在0.5左右时,具有中等影响;当效应值在0.2左右时,影响较小(Borenstein et al., 2009)。本研究计算的合并效应值 $g = 0.683 (p < 0.001)$,介于0.5~0.8间,为中等偏上。因此,总体上教学干预对学生批判性思维能力的提升效果显著。

2. 教学干预对批判性思维技能和倾向的影响

教学干预对学生批判性思维的影响分技能和倾向两个维度。结果显示,倾向和技能维度的效应值分别为0.653($p < 0.001$)及0.732($p < 0.001$),说明教学干预对学生批判性思维的倾向和技能有中等偏上的正向显著影响。异质性检验结果($Q =$

0.034, $p = 0.853$)表明:学生批判性思维倾向和技能之间不存在显著差异。这说明教学干预对学生批判性思维倾向和技能的影响效果是同等的。

(四)调节变量检验结果

1. 不同学科中批判性思维教学有效性的差异

不同学科中教学干预对批判性思维提升效果的差异分析见表一。从结果可以看出,目前绝大多数学科采取的教学干预都是有效的。不同学科中教学干预对提升学生批判性思维的效果见表三,各学科的最终效应值最高的三位依次是数学($g = 1.069, p < 0.05$)、科学($g = 1.062, p < 0.001$)和物理($g = 0.959, p < 0.001$)。从整体来看,各学科的教学干预对批判性思维作用的效应值大部分在0.5以上,说明大部分学科的教学干预都是有效的,且效果值偏高。也有个别学科未达到统计意义上的显著,如生物。组间效应值结果($\text{Chi}^2 = 19.788, p = 0.019$)显示,不同学科的教学干预对学生批判性思维存在显著差异。

研究者进一步对学科再次进行编码,将数学、计算机、生物、科学、综合等编为理科,英语、教育学等编为文科,其他学科按照其教学内容进行编码。文理科中批判性思维教学有效性差异见表二。结果显示,文科、理科之间不存在显著差异,但是在理科中开展批

判性思维教学,效果稍优于文科。

2. 不同学段的批判性思维教学有效性差异

为了检验不同学段学生批判性思维教学效果差异,本研究对四个学段进行了效应值计算。不同学段的批判性思维教学效果的差异分析见表三。结果显示,所有学段的效应值都大于0.5,且都达到了显著性水平。这说明不同学段的教学干预对批判性思维提升具有正向显著的影响。具体来看,高中阶段的教学干预的效应值最高,为1.273($p < 0.001$),超过1,说明教学干预在高中阶段对学生批判性思维具有高度的正向影响;小学、初中阶段的效应值接近0.8,说明教学干预对初中和小学阶段学生的批判性思维提升效果较好;大学阶段的效应值在0.5左右,说明教学干预对大学生批判性思维的提升效果具有中等正向的影响。组间效应值结果显示($\text{Chi}^2 = 10.148, p < 0.05$),不同学段之间存在显著差异,不同学段的批判性思维教学效果有所不同,高中学段的提升效果最强。

3. 不同实验周期批判性思维教学有效性差异

为检验不同实验周期对批判性思维教学有效性的影响,本研究将实验周期分三个时长进行比较。三类实验周期对批判性思维提升效果的差异分析见表

表一 不同学科中批判性思维教学有效性差异

学科	数量	效应值及95%的置信区间				双尾检验		组间效应值 (P值)
		效应值	标准误	下限	上限	Z值	P值	
数学	3	1.069*	0.148	0.779	1.359	7.228	0.023	Chi ² = 19.788 p = 0.019
物理	11	0.959***	0.210	0.548	1.369	4.575	0.000	
生物	6	0.308	0.889	-1.435	2.050	0.346	0.729	
科学	6	1.062***	0.265	0.543	1.581	4.008	0.000	
英语	22	0.823***	0.130	0.568	1.078	6.325	0.000	
教育学	9	0.403**	0.150	0.108	0.698	2.680	0.007	
计算机	2	0.604***	0.142	0.327	0.882	4.795	0.000	
医学	19	0.512*	0.131	0.255	0.769	3.902	0.045	
综合	7	0.854**	0.317	0.233	1.476	2.696	0.007	
其他	18	0.545***	0.095	0.358	0.731	5.712	0.000	

注:***表示p值达到0.001显著水平,**表示p值达到0.01显著水平,*表示p值达到0.05显著水平。

表二 文理科中批判性思维教学有效性差异

学科	数量	效应值及95%的置信区间				双尾检验		组间效应值 (P值)
		效应值	标准误	下限	上限	Z值	P值	
理科	60	0.672***	0.094	0.488	0.856	7.167	0.000	Chi ² = 0.026 p = 0.871
文科	43	0.693***	0.086	0.523	0.862	8.007	0.000	

注:***表示p值达到0.001显著水平。

四。结果显示,所有涉及实验周期的效应值均大于0,说明三种实验周期对学生学习成果都具有显著正向影响。其中,实验周期为“半年以上”的效应值高达0.995($p < 0.001$),说明半年以上的实验周期对学生批判性思维提升产生了高度的正向影响;实验周期为“2个月到半年”的效应值为0.711,表明2个月到半年的实验周期对学生批判性思维的提升产生了中等偏上的正向影响;2个月内的实验周期的效应值为0.455,说明具有中等影响。组间效应值检验结果($\text{Chi}^2 = 6.455, p < 0.05$)显示,不同实验周期对学生批判性思维的教学效果存在显著差异,也就是说,不同实验周期对学生批判性思维的提升效果不同,教学时间越长,学生批判性思维的提升效果越好。

4. 不同实验人数批判性思维教学有效性的差异

本研究将实验人数分为1~30人、31~50人及50人以上三个类别,三个不同实验人数规模的效应值见表五。元分析结果显示,三个规模的实验效应值都在0.5以上,且都达到了显著性水平,说明不同

规模的实验人数对批判性思维具有中等偏上的显著正向影响。组间效应值结果显示($\text{Chi}^2 = 1.625, p = 0.444$),不同实验人数规模不存在显著差异。这说明教学干预的影响效果在不同实验人数上不存在差异,具有同等的影响效果。

5. 技术介入与否对批判性思维教学有效性的差异

本研究将教学干预的类型分为有无技术介入,技术介入与否对批判性思维教学有效性的差异分析见表六。结果显示,效应值都在0.5以上,说明不论有无技术介入,教学干预都对学生批判性思维的提升存在正向显著影响。分析 Chi^2 组间值发现,有无技术介入的教学手段之间对学生批判性思维的提升不存在显著性差异($\text{Chi}^2 = 2.217, p = 0.137$)。

6. 东西方国家批判性思维教学有效性的差异

为了检验教学干预在不同国家(或)地区的适用性,本文将原始数据编码为东西方国家两个类别。东西方国家的批判性思维教学有效性的差异分析见表七。

表三 不同学段中批判性思维教学有效性的差异

学科	数量	效应值及95%的置信区间				双尾检验		组间效应值 (P值)
		效应值	标准误	下限	上限	Z值	P值	
小学	4	0.800***	0.230	0.351	1.250	3.488	0.000	$\text{Chi}^2 = 10.148$ $p = 0.017$
初中	7	0.786***	0.184	0.425	1.147	4.264	0.000	
高中	13	1.273***	0.217	0.849	1.698	5.88	0.000	
大学	79	0.576***	0.073	0.433	0.719	7.919	0.000	

注:***表示p值达到0.001显著水平。

表四 不同实验周期下批判性思维教学有效性的差异

实验周期	数量	效应值及95%的置信区间				双尾检验		组间效应值 (P值)
		效应值	标准误	下限	上限	Z值	P值	
2个月以内	26	0.455***	0.124	0.212	0.697	3.670	0.000	$\text{Chi}^2 = 6.455$ $p = 0.040$
2个月到半年	63	0.711***	0.086	0.543	0.879	8.305	0.000	
半年以上	14	0.995***	0.181	0.640	1.350	5.491	0.000	

注:***表示p值达到0.001显著水平。

表五 不同实验人数批判性思维教学有效性的差异

实验人数	数量	效应值及95%的置信区间				双尾检验		组间效应值 (P值)
		效应值	标准误	下限	上限	Z值	P值	
1-30	30	0.838***	0.154	0.536	1.140	5.434	0.000	$\text{Chi}^2 = 1.625$ $p = 0.444$
31-50	40	0.663***	0.081	0.504	0.822	8.152	0.000	
50以上	33	0.593***	0.117	0.363	0.823	5.062	0.000	

注:***表示p值达到0.001显著水平。

结果显示,东西方国家的效应值分别为0.794($p < 0.001$)和0.419($p < 0.01$),说明东西方国家的批判性思维教学干预对能力提升具有正向影响。尤其是,东方国家的效应值接近0.8,说明教学干预的效果非常好。组间值显示,东西方国家的教学干预对批判性思维的的提升作用存在显著差异($\text{Chi}^2 = 5.367, p = 0.021$)。

7. 不同测评工具对批判性思维教学有效性的差异

为了检验不同测评工具的适用性,本文将测评工具编码为标准与自编测验两类。不同测评工具对批判性思维提升作用的差异分析见表八。结果表明,两种测评工具的效应值都在0.5以上,且都达到了显著性水平,说明两种测验对批判性思维具有正向显著的影响。标准化测评工具效应值为0.561($p < 0.001$),具有中等程度的影响,而自编测验的效应值是1.282($p < 0.001$),具有高度的影响。组间效应值检验结果 $\text{Chi}^2 = 13.795 (p = 0.000)$,说明标准化的测试工具与自编测验之间存在显著差异。

四、结论与讨论

本文采用元分析对79篇实验或准实验研究进行分析,得到如下研究结论:

1) 教学干预对学生批判性思维具有显著提升作用。教学干预对学生批判性思维的倾向和技能都有显著影响。这说明存在有效的教学策略来培养学生的批判性思维。国外的元分析研究也支持本文的结论(Abrami et al., 2015; Niu et al., 2013)。

2) 不同学科领域都有提升批判性思维的有效策略。其中,数学学科效果最好,而生物效果较弱;文理科对批判性思维的影响没有显著差异。部分学科的效应值未达显著性水平,可能与样本数量不足和研究方案的设计有关。

3) 任一教育阶段对批判性思维教育都能起到促进作用,但影响效果在不同学段存在显著差异。分析结果显示,对小学、初中、高中和大学阶段的学生进行教学干预后,其批判性思维能力都有显著提升;高中阶段效果最佳,大学阶段的教学效果最弱。这在一定程度上符合之前一项元分析的研究结果。小学生与中学生无显著性差异,但均显著高于本科生(Abrami et al., 2008)。这一现象说明,批判性思维教学在小学、初中的效果较好,高中阶段则是培养学生批判性思维的最佳时期,这可能与学生的认知发展相关。

4) 不同实验周期对学生批判性思维教学的效果存在显著差异,实验周期越长,批判性思维的提

表六 技术介入与否对批判性思维教学有效性的差异

教学干预	数量	效应值及95%的置信区间				双尾检验		组间效应值 (P值)
		效应值	标准误	下限	上限	Z值	P值	
技术干预	39	0.556***	0.113	0.335	0.777	4.930	0.000	$\text{Chi}^2 = 2.217$ $p = 0.137$
无技术干预	64	0.759***	0.077	0.608	0.910	9.842	0.000	

注:***表示p值达到0.001显著水平。

表七 东西方国家批判性思维教学有效性的差异

地区	数量	效应值及95%的置信区间				双尾检验		组间效应值 (P值)
		效应值	标准误	下限	上限	Z值	P值	
东方国家	70	0.794***	0.068	0.661	0.927	11.726	0.000	$\text{Chi}^2 = 5.367$ $p = 0.021$
西方国家	33	0.419**	0.147	0.132	0.707	2.857	0.004	

注:***表示p值达到0.001显著水平,**表示p值达到0.01显著水平。

表八 不同测评工具对批判性思维教学有效性的差异

衡量工具	数量	效应值及95%的置信区间				双尾检验		组间效应值 (P值)
		效应值	标准误	下限	上限	Z值	P值	
标准化测试	84	0.561***	0.067	0.429	0.692	8.342	0.000	$\text{Chi}^2 = 12.904$ $p = 0.000$
自编测验	19	1.282***	0.189	0.911	1.653	6.772	0.000	

注:***表示p值达到0.001显著水平。

效果越明显。研究结果表明,最有效的干预措施需持续8周以上。这一结果说明,高阶思维能力不能在短时间内获得大幅度提高,只能通过长期的努力培养。这一结果支持了贝哈尔霍伦斯坦(Behar-Horenstein, 2011)的观点。

5) 不同人数对批判性思维的教学有效性不存在显著差异,即不同规模班级的教学效果都差不多。因此,开展具体教学时,无需考虑班级规模的大小。

6) 技术介入与否对批判性思维教学效果的影响不明显。由此可见,以往十年的研究中技术的使用并未对教学效果产生显著影响。产生这一现象的原因可能是学生在技术介入的课程中缺乏充足的时间开展批判性思维和探究性学习。因此,无论有无技术干预,都需要鼓励和激发学生的批判性思维(Arnold et al., 2018)。

7) 东西方国家对批判性思维教学有效性存在显著差异,东方国家的教学效果明显高于西方国家,原因可能是,相较西方国家,东方国家的批判性思维标准化测试起点低,因此有更多的提升空间(Huber et al., 2016)。由此联系到东西方文化的差异,东方国家有“为考试而学习”的共同社会文化,重视学科知识的掌握,而对批判性思维能力的培养重视程度不够;西方国家已开设专门的批判性思维课程,将批判性思维能力的培养作为教育的目标(李晶晶等, 2019)。

8) 不同测评工具对批判性思维教学有效性存在显著差异,自编量表测评的批判性思维的提升效果明显高于标准的测评工具。这一结果支持了安德森等人(Anderson et al., 2001)的观点,使用标准化的批判性思维测量工具通常无法测得批判性思维能力的增长,原因可能是批判性思维的教授大多与特定的学科内容相结合,标准化的测量无法衡量学生在具体领域中使用到的批判性思维技能,故而提升效果不明显;自编量表大多结合特定的学科领域内容,因此更容易测得批判性思维能力的提升效果(Tiruneh et al., 2018)。

五、建议与展望

综合对批判性思维教学效果元分析结果的讨论,本研究提出以下建议:

第一,拓宽批判性思维与学科融合的范围。本研究发现,批判性思维主要应用在医学、英语、物理学

科,极少有学者在文学和艺术(如语文、历史、美术等)学科中教授批判性思维,但不同学科性质对批判性思维的影响不具有差异性,因此,教学者应将批判性思维充分融入各学科,考察学科与批判性思维之间的关系,设计能够提高学生批判性思维的教学方式与模式,打破批判性思维在学科上的壁垒,提高学生的各种核心素养(邓莉等, 2019)。例如,为培养学生的批判性思维,美国中小学的多种学科教材都设有独立且完整的批判性思维专栏及专题活动。

第二,应把批判性思维的教授引入基础教育,并保持教学的连续性与系统性。近十年来,大学阶段的批判性思维研究成果较多。但是,元分析结果表明,批判性思维的提升效果在高中阶段最佳,大学阶段最低。因此,教学人员应积极在基础教育阶段开展教学实验,为批判性思维在基础教育的教授提供理论依据。另外,应将批判性思维的教学融入现有学科中,将其贯穿于整个教育体系。批判性思维属于高阶思维,需要长时间的教授才能产生明显效果。

第三,在批判性思维的教学中,教师不应盲目地追赶技术潮流而摒弃传统的教授方法。首先,教师应根据学习阶段和学习内容选择学习策略,增强学习的有效性(彭正梅等, 2019)。同时,教师选择教学策略要考虑学习环境的结构要素和特征,设计以培养批判性思维为导向的创新性学习环境(戚业国等, 2020)。其次,若将技术引入进教学,教师应考虑学生对技术的接受程度以及该内容是否适用于批判性思维教授的内容,同时也要明确技术提升批判性思维的机理,这样才能够更好地设计将技术融入学科内容的教学策略(毕景刚等, 2020)。再者,教师应该注重培养人的认知与情感的统一,实现学科特色、主体人典型特征与素养教育核心培养目标之间的互动与联系(颜士刚等, 2018)。

第四,使用多元的测评工具测量批判性思维的变化。测评工具是影响学生批判性思维的教学有效性的变量之一。因此,教学者需要关注测评工具的选择,教师可以先采用经典量表和自编量表,但是量表的编制和改编应考虑本土学生的思维过程和习惯,以及语言、文化、成长环境和社会背景等(沈红等, 2019)。除使用经典量表,也应该关注结合学科内容或具体问题情境的测评工具,如游戏测评方式(冷静等, 2020)。

综上,本研究采用元分析方法,通过综合分析已有的研究数据,得出了教学干预对学生批判性思维作用的综合效应值,用量化结果表明了批判性思维是可教的,并发现学科、学段、教学周期、技术介入、东西方文化以及测评工具对学生的批判性思维的教学效果具有一定的调节作用。

本研究局限性在于:由于数据库选取的原因,样本量有限。元分析方法属于探索性分析工具,得出的结果是推断性结论,且受调节变量的影响,因此在使用其结论时仍需谨慎(王翠如等,2018)。另外,围绕批判性思维是否适用于具体学科,恩尼斯主张将教学干预分成四种方式(Ennis,1989):一般方法(general approach)、灌输法、沉浸法以及混合法。之前的元分析中也使用了这种分类来分析哪种教学干预更有效(Abrami et al., 2008; Abrami et al., 2015)。然而,本研究选取的研究样本中,很多研究并没有使用恩尼斯(Ennis,1989)的分类方法,因此本研究没能对教学干预进行详细划分。

[参考文献]

- [1] Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis[J]. *Review of Educational Research*, 78(4): 1102-1134.
- [2] Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis[J]. *Review of Educational Research*, 85(2): 275-314.
- [3] Anderson, T., Howe, C., Soden, R., Halliday, J., & Low, J. (2001). Peer interaction and the learning of critical thinking skills in further education students[J]. *Instructional Science*, 29(1): 1-32.
- [4] Akyüz, H. I., & Samsa, S. (2009). The effects of blended learning environment on the critical thinking skills of students[J]. *Procedia - Social and Behavioral Sciences*, 1(1): 1744-1748.
- [5] Arnold, D. M., Mortensen, C. J., Thoron, A. C., Miot, J. K., Miller-cushon, E. K. (2018). Identifying the optimal course delivery platform in an undergraduate animal behavior research course[J]. *Translational Animal Science*, 2(3): 311-318.
- [6] Behar-Horenstein, L. S., & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature[J]. *Journal of College Teaching & Learning*, 8(2): 25-42.
- [7] 毕景刚,韩颖,董玉琦.(2020). 技术促进学生批判性思维发展教学机理的实践探究[J]. *中国远程教育*, 41(7): 41-49 + 76-77.
- [8] Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-analysis* [M]. Halstead; John Wiley and Sons: 21-32.
- [9] Bowman, N. A. (2012). Effect sizes and statistical methods for meta-analysis in higher education[J]. *Research in Higher Education*, 53(3): 375-382.
- [10] 褚宏启(2016). 核心素养的国际视野与中国立场-21世纪中国的国民素质提升与教育目标转型[J]. *教育研究*, 37(11): 8-18.
- [11] Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.) [M]. Hillsdale, NJ: Lawrence Erlbaum Associates: 18-74.
- [12] 邓莉,彭正梅(2019). 培养具有全球竞争力的人才:基于全球21世纪技能运动的考察[J]. *湖南师范大学教育科学学报*, 18(2): 88-98.
- [13] Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research[J]. *Educational Researcher*, 18(3): 4-10.
- [14] Halpern, D. F. (2001). Assessing the effectiveness of critical thinking instruction[J]. *The Journal of General Education*, 50(4): 270-286.
- [15] Hedges, L. V., Shymansky, J. A., & Woodworth, G. (2009). *A practical guide to modern methods of meta-analysis* [M]. Washington, DC: National Science Teachers Association: 23-45.
- [16] Huber, C. R., & Kuncel, N. R. (2016). Does college teach critical thinking? A meta-analysis[J]. *Review of Educational Research*, 86(2): 431-468.
- [17] Jonassen, D. H. (1999). Designing constructivist learning environments[A]. REIGELUTH C M. *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory* [M]. New Jersey: Lawrence Erlbaum Associates: 215-239.
- [18] 冷静,郭日发,侯嫣茹,顾小清(2018). 促进大学生批判性思维的在线活动设计研究及可视化分析[J]. *电化教育研究*, 39(10): 75-82.
- [19] 冷静,路晓旭(2020). 题库型游戏评测批判性思维能力研究[J]. *开放教育研究*, 26(1): 82-89.
- [20] 李晶晶,潘苏东(2019). 东西方科学课程标准中批判性思维的比较分析[J]. *教育参考*, (1): 48-55.
- [21] Mcmillan, J. H. (1987). Enhancing college students' critical thinking: A review of studies[J]. *Research in Higher Education*, (26): 3-29.
- [22] Naber, J., & Wyatt, T. H. (2014). The effect of reflective writing interventions on the critical thinking skills and dispositions of baccalaureate nursing students[J]. *Nurse Education Today*, 34(1): 67-72.
- [23] National Center for Excellence in Critical Thinking Instruction (1991). *Defining critical thinking* [R]. Santa Rosa: Sonoma State University.
- [24] Niu, L., Behar-Horenstein, L. S., & Garvan, C. W. (2013). Do instructional interventions influence college students' critical thinking skills? A meta-analysis[J]. *Educational Research Review*, (9): 114-128.
- [25] Pascarella, E. T., & Terenzini, P. T. (2006). *How col-*

lege affects students: a third decade of research[J]. *The Journal of General Education*, 55(5): 160-163.

[26] 彭正梅, 伍绍杨, 邓莉 (2019). 如何培养高阶能力-哈蒂“可见的学习”的视角[J]. *教育研究*, 40(5): 76-85.

[27] 戚业国, 孙秀丽. (2020). 我国普通高中学生批判性思维状况与教育应对[J]. *教师教育研究*, 32(2): 63-70.

[28] Reinstein, A., & Lander, G. H. (2008). Developing critical thinking in college programs[J]. *Research in Higher Education Journal*, (1): 78-94.

[29] 沈红, 汪洋, 张青根 (2019). 我国高校本科生批判性思维能力测评工具的研制与检测[J]. *高等教育研究*, 40(10): 65-74.

[30] Tiruneh, D. T., De Cock, M., Gu, X., & Elen, J. (2018). Systematic design of domain-specific instruction on near and far transfer of critical thinking skills[J]. *International Journal of Educational Research*, (87): 1-11.

[31] 王翠如, 胡永斌 (2018). 翻转课堂真的能提升学习成绩吗? 基于38项实验和准实验研究的元分析[J]. *开放教育研究*, 24(4): 72-80.

[32] Wang, X. M., Hwang, G. J., Liang, Z. Y., & Wang, H. Y. (2017). Enhancing students' computer programming performances, critical thinking awareness and attitudes towards programming—An online peer-assessment attempt[J]. *Journal of Educational Technology & Society*, 20(4): 58-68.

[33] 颜士刚, 冯友梅, 李艺 (2018). 素养教育如何落地: 一种面向实践的素养生成机制解释框架[J]. *现代远程教育研究*, (6): 21-27 + 36.

[34] 周榕, 李世瑾 (2019). STEM教学能提高学生学习创造力? ——基于42项实验研究的元分析[J]. *开放教育研究*, 25(3): 60-71.

(编辑: 赵晓丽)

Can Critical Thinking Really Be Taught? A Meta-Analysis Based on 79 Experimental and Quasi-Experimental Studies

LENG Jing & LU Xiaoxu

(Department of educational information technology, Faculty of Education, East China Normal University, Shanghai, 200062, China)

Abstract: *Critical thinking has become one of the important abilities necessary for talents in the 21st century. However, there has been a debate in academia on whether critical thinking can really be taught. In this study, 79 empirical research studies at home and abroad are selected, and a meta-analysis is adopted to explore the teachability of critical thinking. The results show that there are effective teaching strategies to improve students' critical thinking (including skills and dispositions) in all subject areas and at all educational levels. This study explores the differences in the effectiveness in the teaching effects of critical thinking from the aspects of discipline, learning period, experimental period, sample size, technical intervention, national region and assessment tool. It is worth noting that: 1) there are significant differences in the teaching effect of critical thinking among different disciplines, and it works best in mathematics; 2) the teaching of critical thinking works best in high school; 3) The longer the experimental period of teaching critical thinking, the better the improvement will be; 4) Teaching intervention has the same effect on different sample size; 5) There is no significant difference in the improvement of students' critical thinking with or without technical intervention; 6) There are significant differences in the effectiveness of critical thinking teaching between eastern and western countries; 7) Compared with standardized tests, the improvement of critical thinking ability measured by self-designed tools was more significant. This study uses meta-analysis to quantitatively synthesize the empirical research on critical thinking in the past ten years, expecting to provide some enlightenment for the future development of critical thinking.*

Key word: *critical thinking; meta-analysis; instruction intervention*