

基于元分析的在线学习效果评价研究

马 燕¹ 郭 惠 芬² 邓 丽 姣²

(1. 重庆师范大学 智慧教育研究院, 2. 重庆师范大学 计算机与信息科学学院; 重庆 401331)

摘 要:在线学习打破了传统教学模式,使学习者、教师和在线学习资源之间实现跨越时空的多元、智能和深度交互。为了探究在线学习对学生学习效果影响的显著程度和不同对象在线学习效果的差异,研究采用元分析方法对35个研究进行异质性检验,判断研究异质性;运用次群体分析法调节异质性;进行敏感度分析,检验文献的发表年份对研究结果的影响;利用出版偏误检验筛查文献缺失值。研究发现,在线学习合并效应值显著高于传统面对面学习合并效应值。以上结果说明:研究对象中在线学习的干预对学生学习效果具有正向促进作用;针对不同的学习对象,学习效果有显著差异;对不同学科学生学习效果的影响显著程度也不同;研究中的干预周期与效应值呈非线性关系。

关键词:元分析;在线学习;学习效果;异质性检验

中图分类号:I24

文献标识码:A

文章编号:1673-0429(2021)03-0102-11

doi:10.19742/j.cnki.50-1164/C.210311

2018年教育部关于印发《教育信息化2.0行动计划》,要求积极推进“互联网+教育”,加快实现教育信息化。2019年中共中央、国务院印发《中国教育现代化2035》,提出积极推进大学生在线学习,再次强调了教育信息化的重要性,尤其在疫情期间,在线学习成为学生学习的主要方式^[1]。在此背景下,研究在线学习的学习效果显得尤为重要。Ch^[2]等人通过在线学习提升学生学习体验,不仅提供了可访问性、连通性、灵活性的学习机会,还提升学习者之间互动的能力。但Clark^[3]提出使用任何特定的媒介对学习没有任何好处,这一结论引起学术界的热议及对Online Learning的重新思考。不同学者之间的研究结果相互矛盾,不利于推进对在线学习效果的研究,所以需要阅读、梳理相关文献,运用元分析方法对不同研究的结果进行定量分析。采用元分析方法的原因如下:(1)元分析可以发现干预研究中的个体

收稿日期:2020-12-19

作者简介:马 燕(1960—),男,云南昭通人,博士,重庆师范大学智慧教育研究院教授、博士生导师。研究方向:智慧教育与评价、语义网、社会网络分析等。

郭惠芬(1996—),女,云南曲靖人,重庆师范大学计算机与信息科学学院,硕士生。研究方向:元分析、智慧教育及评价。

邓丽姣(1995—),女,云南曲靖人,重庆师范大学计算机与信息科学学院,硕士生。研究方向:教育评价、现代教育实证研究。

基金项目:教育部人文社会科学研究规划基金项目“‘双一流’视野下教师教育学科群及评价体系研究”(18YJA880061);重庆市研究生教育教学改革研究重大项目“基于SNA的教师教育学科群关系网络与评价研究”(yjg201009)。

差异与研究发现之间的重要关系;(2)区分纳入元分析研究的特点独立影响作用的大小,在特定的领域进行干预研究,发现不同研究特征之间的关系。

参考韩啸(2017)^[4]和 Khechine(2016)^[5]的元分析研究,具体步骤如下:(1)根据研究主题检索文献;(2)根据文献筛选标准,排除无用文献;(3)选择文献中符合条件的数据;(4)对数据进行统计分析(描述性统计、随机效果分析、异质性检验、敏感度分析、出版偏误分析);(5)呈现数据分析结果,撰写研究结论。

二、研究述评

(一)在线学习

在线学习是基于互联网的一种学习方式。当前的在线学习包括从学术性讲座到通过网络提供书本上的信息以及基于互联网的课程等。在本综述中,在线学习被定义为部分或全部在互联网上进行的学习。这一定义排除了纯粹基于印刷的函授教育和没有基于互联网的教学组成部分的独立教育软件程序的教育。笔者运用元分析方法研究在线学习的效果,试图回答以下几个问题:与传统学习相比,在线学习的效果如何?以在线学习补充面对面学习是否能促进学习?哪些条件影响在线学习的效果?

通过文献梳理发现,针对在线学习对学生学习效果影响的研究非常少,已有的研究中只有少部分学者设置对照实验评估在线学习对学习效果影响的显著程度;通过梳理变量发现,研究中存在矛盾的结论,一部分研究结果是不显著的。例如 Sexton^[6]在研究中得出,对 4-H 在职人员进行网络培训与传统培训两种结果并无显著差异。O'Dwyer^[7]对代数 I 在线课程的有效性研究中发现学生通过在线学习对自己的代数技能感到不自信。Cooper 和 Smith & Dillon^[8]在一些高等教育的研究中也发现学生对在线课程和传统课程的满意度无显著差异。

另外一部分研究结果是显著的,很多学者认为在线学习有利于提升学生的学习效果。例如,在线学习与传统面对面学习研究中比较有代表性的是 Aberson^[9],其设置实验组和对照组,选取 25 名参与者参加概率统计学课程并且通过传统的方式进行授课;对照组共有 112 名参与者参加统计学入门课程并且参与效果评估。实验结果显示,在课程中使用线上教程对学生学习效果的影响较显著,通过 Fisher 的精确测试, $F(1,22) < 1, P < 0.001$,意味着在课堂上使用线上教程比不使用更有效。在 Zhang 等人^[10]的研究中,将实验环境分为四种,一种是网络环境(组 1 有交互视频、组 2 无交互视频和组 3 无视频),一种是传统的课堂环境。实验结果表明,视频对提高学习效率的价值取决于交互性的提供。组间学习结果(后测)的 P 值,有交互视频($P = 0.005$)、无视频($P = 0.001$)、传统的课堂环境($P = 0.00$)均小于平均差异水平 0.01。与其他环境相比,学生在提供互动视频的线上学习环境中取得了更好的学习成绩和更高的学习满意度,学生在线上学习环境中的学习效果得到显著提升。Bello 等人^[11]通过干预措施让第一组参与者接受传统课程,以在线的形式将材料呈现给第二组,进行了书面测试和实践技能测试并比较课程后分数的差异。第二组的知识增长在书面($P = 0.228$)和实践技能($P = 0.376$)测试方面都略高于第一组,但不显著。学习者满意度的半定量评分显著高于网络组($P = 0.014$),意味着进行线上课件学习是一个有效替代教学和实践的途径。O'Dwyer^[12]等人对路易斯安那州的虚拟学校中代数 I 在线课程的有效性进行了研究,共有 257 名学生参与了该项目。在后测的 25 个项目中,控制组的学生在 18 个项目上的得分超过了比较班。数据还显示,在创建代数表达式的项目上,通过在线课程学习的学生往往比传统学习的学生做得更好。此外 Zwart(2020)^[13]、Bing(2020)^[14]、Ahlam(2020)^[15]等,对在线学习都持积极的态度。

由此可见,目前针对在线学习是否有利于提升学习效果并没有一个系统性的研究。受样本数据影响,大多数学者进行研究时对数据的分析不全面。除此之外,注重学生学习兴趣的在线课堂和在线活动

逐渐增加,对于在线学习提升学生哪方面的学习效果这一问题还没有统一定论,在线学习和传统学习的相对有效性还需要重新讨论。

(二)元分析

元分析是一种定量综合个体实验和先验假设的实证研究,元分析得出的结论往往比元分析中包含的任何研究都更准确可靠。从理论角度来看,元分析作为一种文献综合方法,可以对定量研究结果的统计数据汇总和分析。元分析是定量和描述性文献综述的方法^[16],也是社会和行为科学中最常用的定量综合方法^[17]。从实践的角度来看,元分析常常用于行政科学研究,以确定研究的实证趋势。且元分析能对普遍现象的实证研究进行统计整合^[18],但整合这些信息需要所有研究结果在同一个尺度上。元分析需要对研究的特征及其发现进行测量。重要的特征可以是实质性的,也可以是方法论的。实质性特征通常是针对研究问题的,可能包括干预的类型和长度、受试者的年龄、设置条件、人员培训等。方法论特征是许多元分析所具有的更为普遍的特征,包括样本量、测试可靠性、受试者选择程序、实验设计类型和实验者偏倚程度等因素。对一项研究的实质性和方法论特征进行编码的原因是为了了解研究结果是否因其特征而有所不同。元分析允许研究者为读者提供研究结果的一般描述,以及研究结果在不同研究类型之间变化的信息。对于一个研究不太协调的学科来说,元分析方法达到的效果可能不太理想,但元分析的价值在于能够更好地检查研究结果之间差异的来源,为读者展示研究结果的一般描述和不同研究类型之间的差异。

国内外以元分析为研究方法的文献都聚集在医学领域,教育领域中运用元分析进行的研究较少,而针对在线学习进行元分析的研究更是寥寥无几。其中陈纯瑾通过元分析的方法对比纯粹的在线学习和传统学习与混合学习,研究哪一种学习方法更加有利于提高学生学习效果,但研究对象只有中小學生、本科生和研究生。国外对在线学习进行元分析研究最具有代表性的是 Means,其利用元分析方法对在线学习进行研究,研究对象为 K-12 学生。基于此,本研究扩大研究对象范围,不仅包括中小學生、本科生、研究生,还包括成人大學生、护士、老年人、公司职员,开展在线学习对学习效果影响的综合性定量研究。

三、研究设计

(一)文献搜索

本研究英文文献检索了 Science Direct、ReararchGate、Springer link、Web of Science、Elsevier、ProQuest 等数据库,共检索到 792 篇文献。以 Online Learning、The Impact on Online Learning、A Meta-analysis、Learning Effect 为关键短句,联合 Positive Affect、Negative Affect、Significant Impact、No Significant Effect 为关键词进行检索,并通过 Google Scholar 对文献进行查漏补缺。为了避免出版偏误过大,对检索文献的参考文献进行手工搜索。

(二)文献筛选

结合元分析设定条件和研究主题限制,参考 Means(2010)^[19]对文献的筛选标准,纳入元分析的文献需满足以下标准:(1)使用严格的研究设计;(2)文献是对比在线学习和传统学习的学习效果研究;(3)研究结果之间具有独立性,若两个研究之间的样本相同或存在交叉的情况,只需将报告内容更详细的研究或者大样本研究纳入元分析中;(4)文献必须与学习效果有关,即应具备实验组和对照组;(5)文献中必须有完整的实验结果,即研究提供了实验组与对照组的样本容量 Ne、Nc,均值 Xe、Xc 以及标准差 Se、Sc,或者提供足够的信息来计算效应值(Effects Size);(6)文献中必须包含相关系数(Pearson Coefficient);(7)当一个研究以多种形式发表时(如学位论文、期刊文章),统一选取发表在期刊上的文献用于编码。文献筛选过程如下表:

表 1 文献筛选过程

排除的主要原因	排除数量
1. 文献重复	18
2. 标题和摘要不符合要求	503
3. 文章内容不符合要求	183
4. 没有对学习效果产生影响	12
5. 没有对照组和实验组	14
6. 没有自变量与因变量的相关系数	10
7. 没有提供足够信息	9
8. 其他原因	8

(三)文献编码

文献的研究方法大多涉及结构方程模型,因此选取 35 篇检验数值完整性较好的文章纳入元分析,笔者对其进行全文阅读,提取作者、发表年份、样本量、样本特征、相关系数等内容进行编码,为后续研究提供基础,如表 2 所示:

表 2 文献编码表(部分)

序号	作者(发表时间)	样本量	样本特征	相关系数
1	James S et al. (2009)	264	4-6 年级学生	PRE: r=0.52;POST: r=0.49
2	David L et al. (2009)	77	1 年级学生	PRE: r=0.77;POST: r=0.41
3	Wallace et al. (2000)	93	学士	PRE: r=0.44;POST: r=-0.35
4	Prior et al. (2016)	151	商科硕士	PRE: r=0.616;POST: r=0.543
5	FOX et al. (2007)	200	心理学士	PRE: r=0.71;POST: r=0.71
6	Robert M et al. (2010)	45	学士	PRE: r=0.69;POST: r=0.34
7	Was et al. (2019)	98	学士	PRE: r=-0.33;POST: r=0.36
8	Road et al. (2007)	30	初中生	PRE: r=0.517;POST: r=0.489
9	Tselios N et al. (2001)	108	学士	PRE: r=0.25;POST: r=0.05
10	Chu et al. (2010)	290	成人学生	PRE: r=0.51;POST: r=0.68
11	Law et al. (2010)	133	程序设计学士	PRE: r=0.52;POST: r=0.57
12	Evans et al. (2008)	45	高中生	PRE: r=0.47;POST: r=0.65
13	Beile et al. (2002)	49	硕士	PRE: r=0.56;POST: r=0.49
14	Chu et al. (2010)	317	成人学生	PRE: r=0.59;POST: r=0.75
15	Maki et al. (2002)	134	学士	PRE: r=0.49;POST: r=0.52

注:PRE:干预前;POST:干预后

四、研究设计

(一)方法选定

为了使研究结果更加科学合理,笔者选择随机效应模型,对数据进行异质性检验,了解研究是否有异质性,再对产生异质性的原因进行分析。分析方法包括次群体分析法和元回归分析法,本文选择前者;对数据进行敏感度分析,检验研究的效果量随“时间”的推移是否发生大的变化;对数据进行出版偏误分析,判断元分析结果是否因文章的缺失值而造成偏误,再运用 Trim & Fill 方法进行修正。

(二)效应量计算

元分析中效应量的计算受 Glass(1976)的影响很大。其认为^[20]将研究结果结合起来,首先要计算效应量 g 的估计值,效应量 g 是实验组和对照组平均值之间的标准化差异,如公式(1)所示:

$$g_i = \frac{\bar{Y}E_i - \bar{Y}C_i}{S_i^c}, i=1, \dots, k. \quad [21] \quad \text{公式(1)}$$

其中 $\bar{Y}E_i$ 为实验组(在线学习); $\bar{Y}C_i$ 为对照组(传统学习); S_i^c 为实验组与对照组的标准差,代表了两组之间的差异。不同研究对效应量的估计是标准化的,因此它们在同一尺度上有效。因此,元分析可以在不同研究中合并这些估计值,或者将效应量作为原始数据来分析方差或多元线性回归,从而将研究的特征与学习效果联系起来。

效应量估计值 g 可以分解为一个系统部分,反映真实的群体处理效应;一个非系统部分,反映用于计算效应量的个体得分的抽样误差^[22-24]。系统部分称为种群效应量,如公式(2)所示:

$$\delta = \frac{(\mu^E - \mu^C)}{\sigma} \quad [22-24] \quad \text{公式(2)}$$

其中 μ^E μ^C 分别为实验组和对照组的总体均值, σ 为研究组内的总体标准差。根据 Greek 的定义标准,用 δ μ^E μ^C 和 σ 之间的关系表示种群效应的大小。 δ 是由本研究中观察到的种群参数 μ^E μ^C 和 σ 定义的种群效应量。效应量估计 g 的非系统部分是抽样误差? $=g-\delta$, g 的分解,如公式(3)所示:

$$g = \delta + \varepsilon \quad [22-24] \quad \text{公式(3)}$$

在效应量的统计分析中,必须要将效应量的样本估计分解为系统和非系统部分。元分析中的所有系统关系都涉及到群体效应大小,抽样误差 ε 根据定义是非系统的,因此它与任何要素都没有系统关系。效应量 g 可以用来解释 δ , 因此,如果元分析使用回归或方差分析来研究有效程度与效应量之间的关系,二者的关系就是有效程度与群体效应量之间的关系。

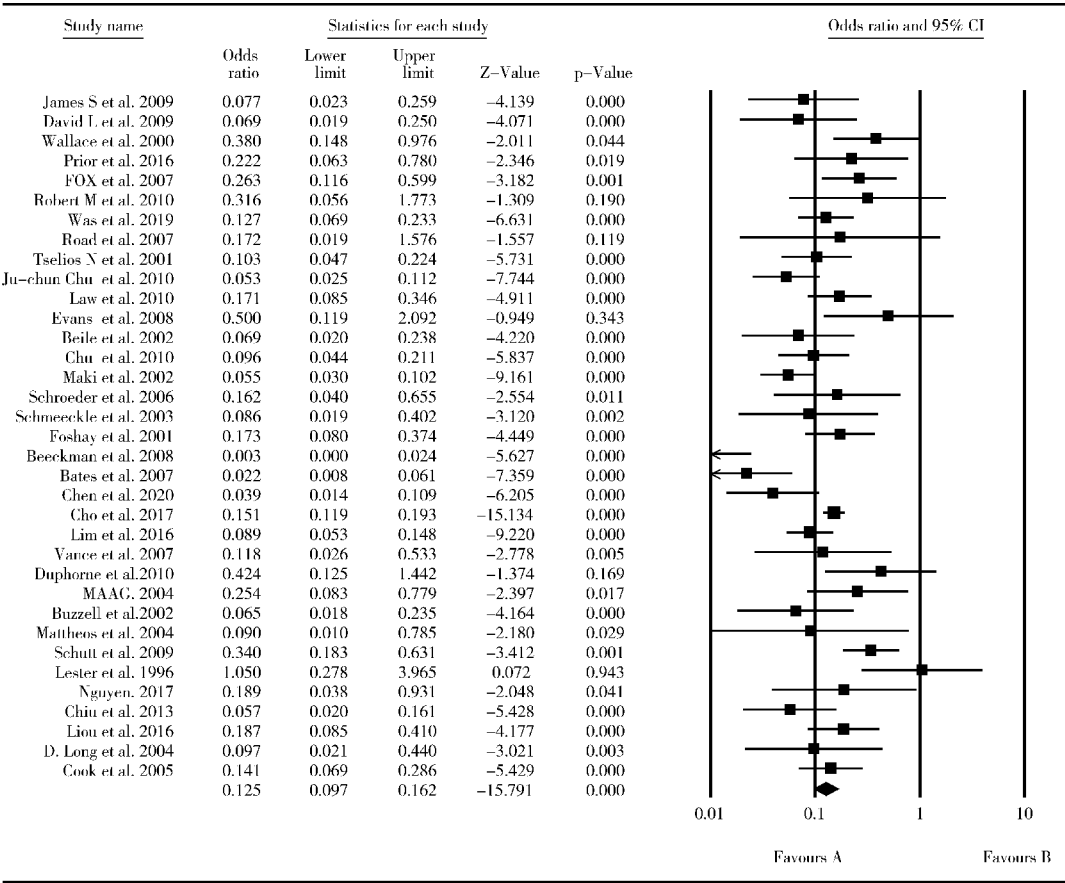
五、数据分析

(一)随机效果模型

随机效果与固定效果的区别在于对权重的修正不同。固定效果对权重的修正是依据样本数作修正,样本数越大,权重越大。但固定效果本身有其缺点,样本数量大的权重非常高,样本数量小的权重非常低,这样会导致忽略小样本对整个研究的影响。而随机效果会将样本数很大的权重缩小,将权重按照一定的比例来分配,将小样本的权重放大,平衡小样本与大样本之间的权重。

固定效应分析对调节变量相关的差异更敏感,但产生错误拒绝原假设的可能性更大。随机效应模型通过在标准误差上增加一个随机常数来减少误差,为了降低测量工具、文化差异、研究对象等对分析结果的影响,笔者采用更加科学合理的随机效果模型。

风险比(Odds Ratio, OR)是实验组发生的比率与控制组没有发生的比率的比值。OR>1 为受到某个条件影响的事件比较多;OR=1 为受某条件影响发生的事件几乎没有;OR<1 受某条件影响发生的事件较少。X 轴为效果量的大小;Y 轴为效果量不显著时的值,当运用的效果量为 OR 时,等于 1 为不显著,运用的效果量为 r 时,等于 0 为不显著。黑色方形越大,权重越高,方形中间对应 X 轴的值为点估计值,也是 Odds Ratio 的值。如图 1 所示,OR 在 0.5-0.9 之间,说明在线环境中发生的事件较多,受到的影响也较大。Lower limit 和 Upper limit 代表 95% 效果量置信区间的左右区间,如果黑线左右区间包含了 1 就代表它不显著。上图中跨越 Y 轴的文献较少,表示整体研究结果显著。底部黑色实心菱形为结合所有研究的平均效果量及置信区间,样本量越大,置信区间就越短,反之亦然。标准误差(Standard Error, SE)是对一组数据可靠性的估计,误差越小,研究结果越可靠,反之则较不可靠。



注:Odds ratio 的计算公式:Log Odds Ratio= $\text{Log}((A * D)/(B * C))$;Log Odds Variance = $(1/A+1/B+1/C+1/D)$;Log Odds SE= $\text{Sqr}(\text{Log Odds Variance})$;Odds Ratio= $\text{Exp}(\text{Log Odds Ratio})$

图 1 随机效果模型

以图 1 中 James S et al(2009)为例进行分析,点估计值为 0.077,左区间为 0.023,右区间为 0.259,且黑线左右区间没有包含 1,所以 Z 值的绝对值 <4.22 ,P 值 <0.01 ,表明这个研究结果显著。Schoenfeld-Tacher、McConnell 和 Graham(2001)^[25]对效应量的估计是+0.80,而图中真实效应量在-0.10 到+1.70 之间的概率为 95%,且 P 值绝大多数都小于 0.01,说明整体的研究结果是显著的。从整体来看,只有 5 个研究是不显著的,其他的研究均是显著的。图中 SE 的数值绝大多数都小于 0.8,说明标准误差较小,研究结果较为可靠,选取的数据也较为可靠。

(二) 异质性检验

表 3 异质性检验

Model		Heterogeneity				Tau-squared		
Model	Q-value	df(Q)	p-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	93.093	34	0.000	63.477	0.315	0.162	0.026	0.561

注:I-squared 计算公式: $I^2 = \text{组间方差} / (\text{组间方差} + \text{组内方差})$;自由度计算公式:Df(Q)= 文献数量-1

异质性检验可以了解研究间的差异是否只是随机误差。产生异质性的原因有很多,例如追踪研究时间长度不同、统计方法不同、调节变量不同等等。当发现研究之间效果量存在异质性时,需要调节变量来解释异质性存在的原因。

如表 3 所示,研究选取 35 篇文献,所以自由度为 34,Q 检验 $p=0.000<0.1$,说明研究结果是显著的且具有异质性。测量异质性对研究结果的影响程度有三个统计量:

$H>2.25$ 表明研究存在高度异质性,计算公式为: $H=\sqrt{\frac{Q}{(k-1)}}$ ^[26-27],如果 $Q<K-1$,则 Q 值会被替换为 0; R 是随机效果的平均标准误差与固定效果标准误差之比; I^2 是 H 的变形,计算公式为: $I^2=\left(\frac{H^2-1}{H^2}\right)$ ^[26-27],它描述了研究中由于异质性造成的事件在总变异中的比重。

异质性 I^2 判断标准为:0% -25% 异质性低;25% -50% 中度异质性;50% -75% 可能有较严重的异质性;75% -100% 需做异质性存在的原因探讨,对造成异质性原因进行分析,方法包括次群体分析法和元回归分析法。但 I^2 在文献不多的情况下精准度很低,所以 I^2 的值在文献数量较多的情况下才具有参考价值。本研究在 34 的自由度下产生的卡方值为 93.093,且 $I^2=63.477$,表明研究有异质性。

T^2 为后续指标计算的基础,本身不是一个检验值,文献越多 T^2 越大, $T^2=0.315$ 说明研究纳入分析的文献数量不多。通过图 1 中的森林图可以看出,黑线超出了 1,代表左右区间包含了 1,显著的研究有 28 个。由于研究存在异质性,需要对数据进行调节变量分析,如果是类别变量就进行次群体分析,如果是连续变量则进行元回归分析。

(三) 调节变量分析

表 4 调节变量分析

Effect size and 95% interval						Heterogeneity		
Group	Number Studies	PointEstimate	Lower limit	Upper limit	Q-value	df(Q)	P-value	I-squared
Fixed effect analysis								
POST	34	0.456	0.437	0.475	736.757	33	0.000	95.521
PRE	34	0.371	0.350	0.391	656.171	33	0.000	94.971
Total within					1392.928	66	0.000	
Total between					36.123	1	0.000	

运用公式($q=Z_A-Z_B$,where? $Z=Fisher's?z$)来计算 Pearson 相关系数 r 的均值,再通过 Fisher's r 到 Z 变换(Zr)来测量研究中相关关系的大小。以干预前和干预后为类别变量,对数据进行次群体分析,如表 4 所示,干预前点估计值为 0.456,左区间为 0.437,右区间为 0.475;干预后点估计值为 0.371,左区间为 0.350,右区间为 0.391。干预前和干预后 P 值为 $0.000<0.05$,说明两个群都是显著的,但两个 I^2 都比较高,说明这两个群是异质的,所以要选择随机效果模型的两群比较数据。两个群的自由度为 1,卡方值差异为 36.123, P 值为 $0.00<0.05$,代表干预前和干预后学生的学习效果有显著差异,并且干预前学习效果没有干预后好。

(四) 敏感度分析

敏感度分析的目的是将某些不合适的论文删除后,看剩余论文的合并效果是否会因此改变,借此测试整合性效果的稳定度。如果移除一篇文章,整合效果改变了,证明这篇文章对整个研究是比较重要的,如果整合效果没有太大改变,证明这篇文章对研究的影响不大。研究运用累积分析法来检验效果量随“时间”变化的效果,了解文献的年份对研究有无影响。由图 2 得知,年份越久远的文献,效果量越接近于 1,越不显著,越接近于现在的文献效果量越不接近于 1,越显著,表明“文献年份”对研究的整体结果有显著影响。

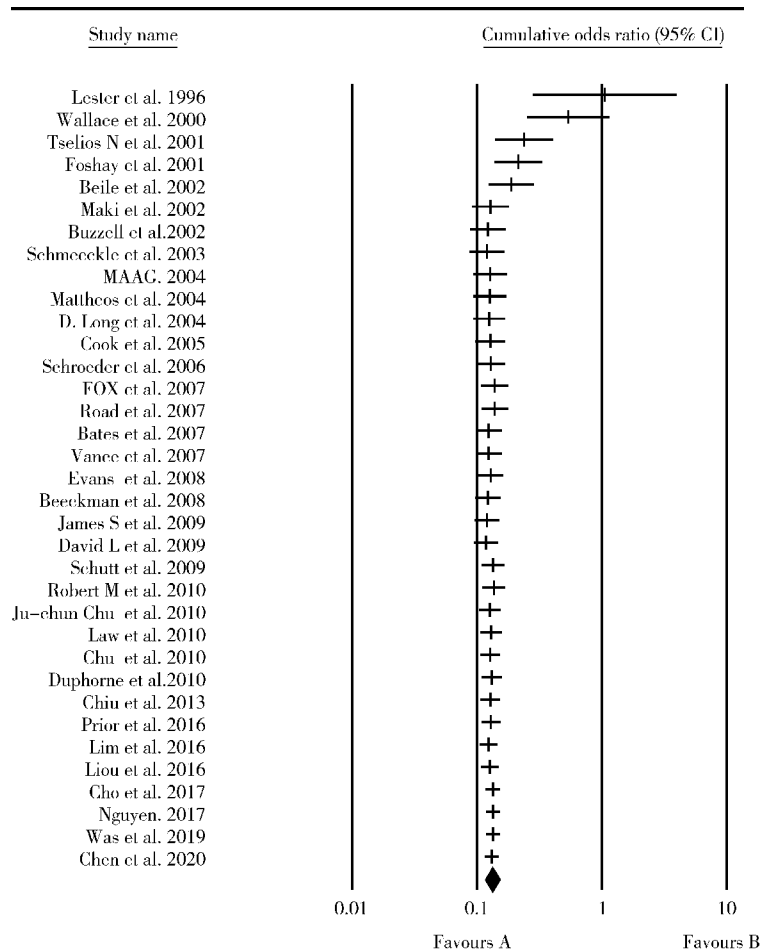


图2 敏感度分析图

(五) 出版偏误检验

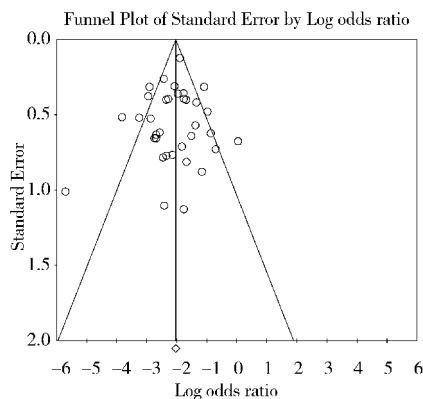


图3 漏斗图

出版偏误可以通过漏斗图(Funnel Plot)或统计分析来鉴定。漏斗图以标准误分之一为纵轴,样本数越大,标准误越小,漏斗图形状越对称。效果大小为横轴,形状类似于倒置的漏斗。如果漏斗图的形状显示非对称,表示存在不显著的文章没有纳入元分析中;如果左右没有平衡或对称,表示纳入的文章存在缺失。并且样本数大的研究分布在上方且接近于平均效果量,样本数小的研究会分布在下方并呈现与平均效果量较分散的位置。如图3所示,每个圆圈代表一篇文献,圆圈都集中在顶部,且以对称的形式均匀分布在左右两侧,说明研究的出版偏误较小。

表 5 Trim&Fill 修正

	Studies Trimmed	Fixed Effects			Random Effects			Q value
		Point Estimate	Lower Limit	Upper Limit	Point Estimate	Lower Limit	Upper Limit	
Observed values		0.13093	0.11443	0.14980	0.12511	0.09666	0.16194	93.09275
Adjusted values	1	0.13317	0.11642	0.15232	0.13151	0.10032	0.17241	106.36677

Trim & Fill 修正方法通过迭代的方式做计算,将剔除的样本补到右边镜设的位置,以此来放大缩小后的效果量方差,最终达到修正出版偏误的效果。如表 5 所示,本研究出版偏误较小,运用 Duval and Tweedie's Trim&Fill 的方法修正了 1 篇文献。通过修正,点估计值从 0.12 修正为 0.13;置信区间从 0.09-0.16 修正为 0.10-0.17;卡方值从 93.09 修正为 106.36。

五、结论

(一)讨论

从纳入元分析的 35 项研究中提取效果量,在线学习合并效应值(0.443)显著高于传统面对面学习合并效应值(0.371),表明在线学习比与之进行比较的传统学习更加有效。然而这种整体效果归因于线上线下混合的学习方法优于纯粹的传统学习方法。Clark^[28]认为媒介不是造成学习效果差异的主要因素,多媒体的使用提升了学习效果,是因为结合了其他条件(学习方法、学习者学习能力、教学设计等)。进行异质性检验可以发现不同的研究之间存在显著差异。这种差异证明了运用调节变量分析方法来解释异质性的合理性,调节变量分析发现前后调节变量显著性 $p < 0.05$,在线学习与传统学习相比较,差异更加显著。

研究中的效果量不因样本数量或设计类型的变化而变化,是因为在随机分配实验设计的研究($p < 0.001$)和样本规模较大的研究($p < 0.01$)中,在线学习比传统学习产生了更好的学习效果。研究的敏感度分析发现 2008 年以前的研究和以后的研究平均效果量有差异,且差异与本研究主题有关。比较各种在线学习的叙述性综合研究,学生进行在线学习时,如果提供优质的在线学习环境,设计促使学生进行反思的环节,以及根据学生的具体需求进行个性化在线教学,对学生在线学习效果有促进作用。

(二)结论

本文将元分析运用到在线学习对学习效果影响程度的领域,拓宽了元分析的研究领域。对 35 篇在线学习对学习效果产生影响的文献进行定量分析并检验;对文献进行整理与记录,提取文献中实验组和控制组的样本总量,进行实验的样本数量与相关变量的数值和变量之间的相关系数进行统计分析;对相关变量进行异质性检验、出版偏误分析和敏感度分析,并针对各个问题提出相应的修正方法,最后通过数据分析得出结论:

1. 综合计算随机效果模型数据显示,在线学习的合并效应值为 0.443,高于传统学习的合并效应值 0.371,说明相较于传统学习,在线学习的干预对学生学习效果具有正向促进的影响。

2. 从整体上看,在线学习比传统的面对面学习更有效,主要是因为在线学习的学习方式趋于多样化,学习时间较为灵活,为学生提供多样性选择等。与单纯的在线教学相比,结合在线和面对面方式的教学具有更大的优势。在具体学科运用时,例如医学、计算机等理、化、生类学科中应用效果最为显著。

3. 通过不同群体效应量分析得出,不同群体的在线学习效果不同。且在不同干预周期中运用时,干预周期与效应值呈非线性关系。在线学习的干预对小学生、中学生、高中生及大学学生的学习成效都有显著影响效果,且随年龄增长干预效果越显著。如果单独考虑学习者的年龄,平均效应量对于本科生和其他年龄较大的学习者显著为正,而对于低年级学生则不显著。

本研究存在两个不足之处:1. 本研究纳入元分析的文献均为外文文献,中文文献没有纳入分析样本中,这对研究应用到国内实践会造成局限;2. 样本容量较小,限制了调节变量的检验效力,实验时长、实验对象、实验环境等因素可能会对研究结果产生干扰。

[参 考 文 献]

- [1] 中共中央、国务院印发《中国教育现代化 2035》[EB/OL]. (2019-02-23) [2020-12-19]. http://www.moe.gov.cn/jyb_xwfb/s6052/moe_838/201902/t20190223_370857.html.
- [2] Ch S K, Popuri S. Impact of online education: A study on online learning platforms and edX[C]//2013 IEEE International Conference in MOOC, Innovation and Technology in Education (MITE). IEEE, 2013: 366-370.
- [3] Clark R E. Reconsidering research on learning from media[J]. Review of educational research, 1983, 53(4): 445-459.
- [4] 韩啸. 整合技术接受模型的荟萃分析: 基于国内 10 年研究文献[J]. 情报杂志, 2017, 36(08): 150-155+174.
- [5] Khechine H, Lakhal S, Ndjambou P. A meta - analysis of the UTAUT model: Eleven years later[J]. Canadian Journal of Administrative ences, 2016, 33(2): 138 - 152.
- [6] Sexton, Julie S, Raven, Matt R, Newman, Michael E. A Comparison of Traditional and World Wide Web MethodPREogies, Computer Anxiety, and Higher Order Thinking Skills in the In-Service Training of Mississippi 4-H Extension Agents. [J]. Journal of Agricultural Education, 2002, 43(3): 25-36.
- [7] O'dwyer L M, Carey R, Kleiman G. A study of the effectiveness of the Louisiana Algebra I online course[J]. Journal of research on technology in education, 2007, 39(3): 289-306.
- [8] Cooper L W. A comparison of online and traditional computer applications classes[J]. The Journal, 2001, 28(8).
- [9] Aberson C L, Berger D E, Healy M R, et al. Evaluation of an interactive tutorial for teaching hypothesis testing concepts [J]. Teaching of Psychology, 2003, 30(1): 75-78.
- [10] Zhang D, Zhou L, Briggs R O, et al. Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness[J]. Information & management, 2006, 43(1): 15-27.
- [11] Bello G, Pennisi M A, Maviglia R, et al. Online vs live methods for teaching difficult airway management to anesthesiology residents[J]. Intensive care medicine, 2005, 31(4): 547-552.
- [12] O'dwyer L M, Carey R, Kleiman G. A study of the effectiveness of the Louisiana Algebra I online course[J]. Journal of research on technology in education, 2007, 39(3): 289-306.
- [13] Zwart D P, Van Luit J E H, Noroozi O, et al. The effects of digital learning material on students' mathematics learning in vocational education[J]. Cogent Education, 2017, 4(1): 1313581.
- [14] Xu B, Chen N S, Chen G. Effects of teacher role on student engagement in WeChat-Based online discussion learning[J]. Computers & Education, 2020, 157: 103956.
- [15] Alghamdi A, Karpinski A C, Lepp A, et al. Online and face-to-face classroom multitasking and academic performance: Moderated mediation with self-efficacy for self-regulated learning and gender[J]. Computers in Human Behavior, 2020, 102: 214-222.
- [16] Rosenthal R, Dimatteo M R. Meta-Analysis: Recent Developments in Quantitative Methods for Literature Reviews[J]. Annual Review of PsychPREogy, 2001, 52(1): 59-82.
- [17] Hedges L V, Olkin I. Statistical methods for meta-analysis[M]. Academic press, 2014. 25-42.
- [18] Glass G V, McGaw B, Esmith M L. Meta-Analysis in Social Research[J]. Exceptional Education Quarterly, 1983, 4(3): 113-116.
- [19] Means B, Toyama Y, Murphy R, et al. Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies[J]. Us Department of Education, 2009.
- [20] Glass G V. Primary, secondary, and meta-analysis of research[J]. Educational researcher, 1976, 5(10): 3-8.
- [21] Cohen J. A power primer[J]. Psychological bulletin, 1992, 112(1): 155.

- [22,23,24] Larry V. Advances in statistical methods for meta-analysis[J]. New Directions for Program Evaluation, 1984, 1984(24):25 - 42.
- [25] Schoenfeld-Tacher R, McConnell S, Graham M. Do no harm—A comparison of the effects of on-line vs. traditional delivery media on a science course[J]. Journal of Science Education and Technology, 2001, 10(3): 257-265.
- [26,27] Higgins J P T, Thompson S G. Quantifying heterogeneity in a meta - analysis[J]. Statistics in medicine, 2002, 21(11): 1539-1558.
- [28] Clark R E. Reconsidering research on learning from media[J]. Review of educational research, 1983, 53(4): 445-459.

Research on Evaluation of Online Learning Effect Based on Meta-Analysis

Ma Yan¹ Guo Huifen² Deng Lijiao²

(1. Wisdom Education Research Institute, Chongqing Normal University, Chongqing 401331;

2. School of Computer and Information Science, Chongqing Normal University, Chongqing 401331, China)

Abstract: Online learning breaks the traditional teaching model and enables learners, teachers and online learning resources to achieve multiple, intelligent and deep interactions across time and space. In order to explore the significant degree of the influence of online learning on students' learning effect and the difference of online learning effect of different objects. In this study, meta-analysis was used to test the heterogeneity of 35 studies to judge the heterogeneity; sub-group analysis was used to adjust heterogeneity; sensitivity analysis was used to test the influence of the year of publication on the results of the study; publication error test was used to screen the missing value of the literature. It is found that the merging effect of online learning (0.443) is significantly higher than that of traditional face-to-face learning (0.371). The above results show that the intervention of online learning has a positive promoting effect on students' learning effect; for different learning objects, the learning effect is significantly different; the impact on the learning effect of students in different disciplines is also different; there is a non-linear relationship between the intervention period and the effect value in the study.

Keywords: meta-analysis; online learning; learning effect; heterogeneity test

[责任编辑:刘力]