

多媒体学习中表征认知负荷的眼动指标研究

王国华¹, 田梁浩¹, 聂胜欣¹, 朱珂², 梁云真¹

(1.河南师范大学 教育学部, 河南 新乡 453000;

2.智能教育河南省协同创新中心, 河南 新乡 453000)

[摘 要] 通过眼动设备进行认知负荷的测量已成为认知负荷测量的重要方法, 尽管相当多的研究依赖眼动仪输出的测量数据进行认知负荷表征, 但究竟哪些眼动数据能够准确表征学习者认知负荷状态并没有明确的指标。研究采用循证研究的元分析方法, 对 62 篇国内外实验研究文献进行量化分析, 系统检验认知负荷对眼动各项指标的影响, 探究可以表征多媒体学习中认知负荷状态的眼动指标。研究发现: (1) 纳入研究的总效应值为 0.547 ($k=318$), 表明整体上眼动指标受认知负荷影响较大; (2) 将不同测量指标作为调节变量分析, 发现瞳孔直径等 15 项具有较大效应值的眼动指标能够较好地表征学习者的认知负荷状态; (3) 在其他调节变量分析中, 发现低于 60Hz 的采样率更适用于多媒体学习中认知负荷的测量, 而知识性质和认知负荷来源不存在调节效应。

[关键词] 认知负荷; 多媒体学习; 眼动实验; 眼动追踪技术; 元分析

[中图分类号] G434

[文献标志码] A

[作者简介] 王国华 (1988—), 男, 河南濮阳人。讲师, 博士, 主要从事在线学习及认知状态评估研究。E-mail: wgh19892008@126.com。

一、引 言

认知负荷是描述学习者在处理具体任务时认知系统负荷量的多维结构要素^[1]。过高的认知负荷会导致学习者无法有效处理信息, 从而影响学习效果。而适度的认知负荷可以激发学习者的学习兴趣和动力, 促进学习的深度和广度。认知负荷表征任务需求的方式与身体负荷表征能量需求的方式类似, 需要消耗大脑的信息处理能力, 所以很多研究者使用生理测量的方法来测量人脑的认知负荷^[2]。生理测量法通过评估实验对象进行任务时产生的生理指标变化来测量认知负荷^[3]。

而在认知负荷研究中, 非侵入式的眼动仪器也被广泛应用于认知负荷的测量^[4], 但相关研究存在一些问题, 首先, 研究结论存在一定差异性, 究竟哪些指标能够有效表征多媒体学习中学习者的认知负荷? 这是一个亟待解答的问题; 其次, 大多数研究者是在特定

因素、技术条件下探讨认知负荷的测量, 研究结论适用场景单一, 缺少共性的研究结论。因此, 本研究将使用元分析法 (Meta-analysis) 有目的地整合国内外使用眼动技术测量认知负荷的实验与准实验研究结果, 力图从整体上厘清眼动测量指标体系及其调节变量。

二、文献综述

(一) 眼动指标相关研究

眨眼 (Blink) 相关指标是眼动追踪实验中的常用指标, 包括眨眼次数、眨眼频率、眨眼时间和眨眼间隔。Rosenfield 等人在探究眨眼率与认知负荷之间的关系时发现, 在认知负荷高的任务中, 实验对象的眨眼率显著降低^[5]。但是在 Benedetto 等人的研究中, 15 个实验对象在进行单一驾驶任务和双重驾驶任务时的眨眼频率并没有显著差异^[6]。

眼跳 (Saccade) 的常用指标包括眼跳幅度、眼跳

基金项目: 2021 年度教育部人文社会科学研究青年项目“多模态生理数据驱动的在线学习认知负荷测评模型及方法研究”(项目编号: 21YJC880072); 2023 年河南省教师教育课程改革研究重点项目“职前教师智能教育素养培养模式构建与实践应用”(项目编号: A010)

峰值速度、眼跳距离、眼跳速度、眼跳范围、返回眼跳、眼跳时长、眼跳次数、眼跳频率等。近来的研究表明,眼跳运动能够高效地反映人的信息处理能力,能够连续不断地接受刺激并作出强有力的反应,是一种自发的适应性眼动行为,学习者能够快速调整视觉系统对刺激的处理和注意焦点,因此,学习者认知负荷越高,眼跳的频率也会越高^[7]。Salvucci 等人研究发现,学习者在双任务和主任务执行过程中眼跳频率并没有显著差异^[8]。

瞳孔(Pupil)的眼动指标包括瞳孔的直径或大小、瞳孔扩张、瞳孔面积标准差等。瞳孔大小之所以发生变化,是因为认知负荷的增加致使大脑需要更多的认知资源来完成任务,人就会不自主地控制瞳孔扩张以增加进入视网膜的光线量来提高视觉清晰度^[9]。Krzysztof 等人的研究也发现,通过观察瞳孔大小在实验过程中的变化能够区分不同水平的认知负荷^[10]。

注视(Fixation)的眼动指标包括注视次数、注视点个数、注视持续时间、总注视时间、注视时间占比、注视频率等。Marquart 等人的研究表明,随着任务难度的增加,实验对象的注视持续时间、总注视时间和注视次数均显著增加,而注视点个数和注视时间占比则显著减少^[11]。但是在 Savage 的研究中,17 位学生解答两个不同难度水平的谜题时所测得的注视持续时间却没有显著差异^[12]。

另外,兴趣区(AOI)分析虽然不是严格意义上的度量标准,但其定义了计算眼动指标度量标准的区域。这类分析是指选择一个或多个显示的刺激目标区域,并且专门在该区域提取一些眼动数据来进行统计学分析^[13]。例如:选择学习材料中的文本部分与图片部分,分别统计实验对象注视或停留在每个区域的时间、次数等视觉加工信息,最终就能够得到关于该 AOI 的眼动指标。有研究表明,通过比较不同认知负荷条件下的 AOI 眼动指标,可以深入探究不同认知负荷水平下实验对象的视觉加工特征和认知策略,进而揭示认知负荷对视觉加工和认知加工的影响机制^[14]。然而,在 Johnson 等人的研究中,虽然 AOI 两区转换次数最终在两种不同学习材料设计方式的图文学习任务中呈现出显著不同,但 AOI 注视比例和 AOI 注视时间却没有显著差别^[15]。

可见,虽然使用眼动指标测量认知负荷的研究数量众多,但是究竟哪些眼动指标可以精准表征学习者的认知负荷状态?这一问题存在一定的争议,依然没有明确答案。因此,本研究使用元分析方法对国内外相关量化研究进行梳理与分析,以期找出能够精确表征学习者认知负荷的指标集合,为相关研究提供借鉴。

(二)其他影响认知负荷测量的因素

有学者提出,程序性知识或认知技能的获得与认知负荷存在直接的关系^[16]。这是因为在关于学习者认知图式的系统模型中,程序性知识的获取和使用相比陈述性知识会更加充分地调用工作记忆,工作记忆中被触发的基本单元的数量、工作循环的次数、条件匹配的数量等又恰恰表征了认知负荷的大小^[17]。因此,相较于学习陈述性知识,学习者在学习程序性知识时需要进行更多的认知加工和更深入的思考,其认知负荷很有可能更加明显地通过眼睛的生理行为表现出来。

从采样所得的最直观数据来看,眼动测量指标的基本单位大多与频率相关。这说明眼动仪采样的频率很有可能从统计层面直接影响到测量指标的准确性。眼动仪的采样频率越高,获得的眼动行为信息就越丰富^[18]。高分辨率的眼动仪器能够捕捉到微小的眼动行为,但由于人眼部神经和肌肉的复杂性,某些微小的眼动行为可能并不是认知负荷测量所需的数据。因此,在测量认知负荷的实验中,眼动仪采样率并不一定越高越好。区分眼动仪采样率高低的频率界限为 60Hz,因为在采样技术进步之前,许多屏幕式和穿戴式眼动仪都以这种采样速度运行,并且 60Hz 一直是摄像技术中最常见的频率^[19]。

此外,其他影响认知负荷的因素也是多种多样的,已有相关实验的因变量大致可以分为三类:第一,任务的难度和复杂度不同,所要求的认知资源和信息处理能力就不同,带来的认知负荷也不同;第二,个体的认知能力、经验或知识水平不同,执行任务所采用的认知策略和认知图式自动化程度也会不同,对认知负荷就会产生影响;第三,学习材料组织结构和呈现方式等因素同样会对多媒体学习环境中的认知负荷产生影响^[20]。这三种认知负荷来源经常在实验中被用作控制认知负荷水平的因变量,因此,实际研究中要特别注意认知负荷来源是否会对认知负荷在眼动数据上的表现造成影响。

综上所述,本研究拟采用元分析方法系统梳理国内外已有的实验和准实验研究结果,探讨认知负荷对眼动设备各项指标的影响,主要拟解决以下两个研究问题:(1) 哪些眼动测量指标可以表征多媒体学习中的认知负荷;(2) 在使用眼动指标进行认知负荷测量时还受到哪些变量的调节作用?

三、研究设计

元分析常用的效应指标包括相关系数(Correlation, 以下简称 r)、均数差(Mean Difference)、加权均数差

(WMD)和标准化均数差(SMD)等^[21]。根据 Rosenthal 和 DiMatteo 的研究,用乘积矩相关性系数 r 作为效应值比 Cohen's d 或 Hedges' g 等效应值更能从实际重要性角度解释研究结果^[22]。因此,本研究将采用相关性系数 r 作为效应值以表征认知负荷对眼动仪各项测量指标的影响大小,使用元分析软件(CMA 3.0)进行数据分析,同时运用 Excel 2019 作为编码工具。根据元分析指导准则(PRISMA)的相关步骤及内容开展元分析研究过程^[23],主要包括以下步骤:

(一)文献检索

为了保证原始文献的质量与数量,外文文献的检索选择 Web of Science、EBSCO、Science Direct 等权威英文数据库,组合逻辑检索语句为 TS=(workload OR processing load OR cognitive workload OR mental workload OR physiological OR measurements) AND TS=(eye OR pupil OR blink OR fixation OR pupillary response OR pupillometry OR eye movement),再借助谷歌学术的文献引文进行检索补充。中文文献通过中国知网数据库搜索核心期刊、CSSCI 期刊和学位论文,以“认知负荷(主题)AND 眼动 + 快速眼动 + 眼球运动 + 注视眼动 + 瞳孔 + 眼跳 + 扫视 (关键词)AND 多媒体学习 + 多媒体教学 + 计算机辅助学习 + 电子学习 + 在线学习 + 虚拟学习(关键词)”为检索式。同时使用“滚雪球”的形式进行相关文献的纳入,中英文文献发表时间均限定为 1997 年 11 月—2022 年 11 月,一共得到 286 个检索结果。将所有文献导入 EndNote 文献管理软件,去除重复文献后,共得到文献 256 篇。

(二)文献筛选及纳入标准

本研究文献纳入标准如下:第一,主题必须包括多媒体学习领域认知负荷的测量;第二,研究方法应是实验研究或准实验研究;第三,自变量和因变量分别为认知负荷和测量指标;第四,论文提供了完整的计算效应值 r 的数据,如样本量、方差分析相关值、 p 值或直接

提供相关系数 r 等。各研究之间样本相互独立,若样本重复,则选取更详细或更大样本的研究。经过筛选最终得到 24 篇中文文献和 38 篇英文文献,因大部分文献包含多个眼动测量指标的数据,最终得到 62 篇文献共 318 条数据,其中,实验样本共 14045 人。

(三)文献编码

提取 62 篇文献的信息并进行编码,编码内容涵盖作者+出版年份、眼动指标、知识性质(陈述性知识/程序性知识)、认知负荷来源(任务难易程度/学习材料设计方式/学生学业水平)、眼动仪采样率(60Hz 以下/60Hz 及以上)、总样本量、组间自由度、组内自由度、计算所得的效应值 r ,以及标准化后的 Z_r 。研究涉及的学段和学科涵盖基础教育到高等教育的大多数学科,任务类型包含阅读、搜索、解题、计算机模拟实验等。表 1 为文献信息的部分编码汇总表。

(四)数据处理与分析方法

由于影响认知负荷测量结果因素的多元性和复杂性,以及不同研究所处的背景和地域存在较大差异,所以使用随机效应模型(Random Effects Model,简称 REM)计算 318 个效应值的综合效应值。又因效应值 r 很可能呈现非正态分布,所以为了合并来自多项研究的效应值,首先需要通过 Fisher 的 r -to- Z_r 变换,将效应值 r 转换为 Z_r 值来标准化每条数据的效应值^[28]。转换之后根据各种调节变量合并研究结果,最后通过 Fisher 的 r -to- Z_r 逆向变换计算效应值结果和其估计的 95%置信区间。

将 318 条眼动数据录入 CMA3.0 软件中运行分析得出,基于卡方检验进行的异质性分析的结果非常显著($Q=2987.08, I^2=89.39\%$),卡方检验的自由度 $df=317$,卡方临界值 $p<0.05$ 。根据 Higgins 界定异质性高低的标准, $I^2>75\%$,该结果说明这些通过方差分析得到的效应值变异太大,不能仅用抽样误差来解释^[29]。因此,需要进行调节变量分析以识别导致研究之间异

表 1 特征值编码汇总表(部分)

作者(年份)	眼动指标	知识性质	采样率	认知负荷来源	总样本量(N=14045)	r 值	Z_r
叶爱敏(2016) ^[24]	注视次数	陈述性	60	任务难易程度	40	0.3431	0.3576
	瞳孔大小	陈述性	60	任务难易程度	40	0.1923	0.1947
	眼跳次数	陈述性	60	任务难易程度	40	0.3854	0.4064
Rop(2018) ^[25]	总测试时间	陈述性	250	学生学业水平	75	0.6034	0.6985
Tokuda(2011) ^[26]	返回眼跳	陈述性	50	任务难易程度	16	0.8258	1.1748
	瞳孔大小变化	陈述性	50	任务难易程度	16	0.7362	0.9421
Yang(2012) ^[27]	注视次数	程序性	60	任务难易程度	41	0.4207	0.4485
	眨眼频率	程序性	60	任务难易程度	41	0.1183	0.1189
	眼跳次数	程序性	60	任务难易程度	41	0.2757	0.2830

质性的因素。

(五)发表偏倚检验

发表偏倚(Publication Bias)是指已发表的研究文献不能完全代表实际研究总体的情况,从而引起的结果偏差^[30]。在元分析中,如果出现较大水平的发表偏差,将会威胁到结果的有效性,因此,有必要对原始文献进行发表偏倚检验。本研究使用了定性的漏斗图方法(Funnel Pot)和定量的 Egger 回归法来检验发表偏差。当所筛选的原始文献没有较大的发表偏差时,基于原始文献效应值绘制的漏斗图将呈现为左右较为对称的倒立型漏斗形状,否则将出现明显的左右不对称。本研究所选的原始文献效应值形成的漏斗图以 $Z=0.540$ 为对称轴,左右均匀分布,表明所选的原始文献发表偏差可以接受,如图 1 所示。同时,Egger 回归法的检测结果显示, $t=0.730,p=0.233>0.05$,表明所选的原始文献发表偏差不显著,本研究的结果具有稳健性。

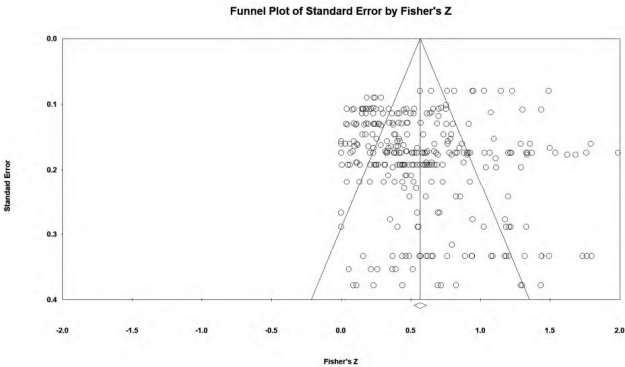


图 1 发表偏倚检验漏斗图

四、研究结果和分析

(一)整体效应分析

进行综合效应值分析时,采用未加权的效应值进行分析将在很大程度上减少个别样本量过大且变异性水平较高的研究对结果造成的影响($I^2=19.87<25\%$),最终得到元分析的随机效应模型的综合效应值 $r=0.547$ ($CI=[0.507,0.584]$)。依据 Cohen 提出的标准,效应值 r 在 0.10~0.30 之间被视为小效应值,0.30~0.50 之间被视为中等效应值,而大于 0.50 则被视为大效应值^[31]。通过综合检查元分析中包含的所有研究,我们认为,认知负荷和眼动测量之间存在较高水平的相关关系,眼动仪各项指标受认知负荷的综合影响较为显著。

(二)测量指标的效应分析

通过计算得到各文献中每一项眼动指标的未加权随机效应模型的效应值和置信区间,具体数据见表 2。将得出的效应值 r 与已知的基准比较,可以确定这些眼动指标受认知负荷的影响大小^[31]。

表 2 以测量指标作为调节变量的分析

	眼动指标	研究数量	效应值	95%置信区间
眨眼相关指标	眨眼时间	6	0.424	[0.089,0.673]
	眨眼频率	10	0.359	[0.146,0.541]
	眨眼次数	4	0.357	[0.260,0.446]
	眨眼时长标准差	1	0.552	—
	眨眼间隔	1	0.677	—
眼跳相关指标	眼跳幅度	10	0.535	[0.249,0.736]
	眼跳峰值速度	3	0.256	[-0.358,0.715]
	眼跳距离	3	0.872	[-0.390,0.996]
	眼跳速度	2	0.804	[-1.000,1.000]
	眼跳范围	1	0.255	—
	返回眼跳	5	0.481	[0.000,0.781]
	眼跳时长	3	0.787	[-0.372,0.987]
	眼跳次数	3	0.527	[-0.193,0.878]
	眼跳频率	1	0	—
瞳孔相关指标	瞳孔大小	20	0.484	[0.329,0.614]
	瞳孔面积标准差	8	0.558	[0.106,0.819]
	瞳孔大小变化	14	0.759	[0.615,0.854]
注视相关指标	注视次数	40	0.598	[0.507,0.676]
	注视点个数	11	0.52	[0.289,0.694]
	注视持续时间	49	0.484	[0.352,0.596]
	总注视时间	39	0.561	[0.416,0.678]
	注视时间占比	24	0.491	[0.359,0.604]
	注视点注视时长标准差	1	0.399	—
	注视次数占比	3	0.475	[-0.491,0.917]
	注视频率	3	0.606	[-0.511,0.962]
	水平注视位置	2	0.491	[-0.763,0.969]
	垂直注视位置	2	0.783	[-0.982,1]
	注视点水平距离	1	0.608	—
	注视点垂直距离	1	0	—
AOI 相关指标	首次注视 AOI 的时间	5	0.401	[-0.171,0.771]
	视线首次到达 AOI 的时间	2	0.946	[-1.000,1.000]
	AOI 进出次数	10	0.391	[0.249,0.517]
	AOI 停留时间	6	0.648	[0.035,0.907]
	AOI 访问次数	5	0.519	[0.355,0.653]
	AOI 平均访问时间	3	0.451	[-0.425,0.891]
其他	AOI 总访问时间	1	0.259	—
	视野范围	1	0.662	—
	总测试时间	13	0.644	[0.483,0.763]
	视觉熵值	1	0.229	—
	总计	318	0.547	[0.507,0.584]

效应值水平最高的测量指标是“瞳孔大小变化”,效应值 $r=0.759,CI=[0.615,0.854]$,表明该测量指标的数值受认知负荷影响最显著。然而就数值大小而言,“眼

跳距离”的效应值似乎更高($r=0.872$),但由于该效应值估计的 95%置信区间包括 0($CI=[-0.390,0.996]$),所以不具有统计学意义;同样,即使“眨眼间隔”的效应值水平很高($r=0.677$),但支撑该效应值的样本量太小($k=1$),结果稳健性不能得到保证,无法确定其是否能够有效表征学习者的认知负荷。

除“眼跳距离”之外,还有一些测量指标估计的 95%置信区间包含 0。这些测量指标包括“眼跳峰值速度”“眼跳速度”“返回眼跳”“眼跳时长”“眼跳次数”“注视频率”“水平注视位置”“垂直注视位置”“注视点垂直距离”“首次注视 AOI 的时间”“视线首次到达 AOI 的时间”“AOI 访问时间”,说明这些测量指标受认知负荷影响不够显著。除“眨眼间隔”之外,还有一些测量指标也因为样本量不足需要进一步探索,如“眨眼时长标准差”“眼跳范围”“眼跳频率”“注视点注视时长标准差”“注视点水平距离”“注视点垂直距离”“AOI 总访问时间”“视野范围和视觉熵值”。以上两类指标被排除在有效测量指标之外。

将其余受认知负荷影响显著的测量指标按照效应值由大到小排序,处于大效应值水平的依次是“瞳孔大小变化”“AOI 停留时间”“总测试时间”“注视次数”“总注视时间”“瞳孔面积标准差”“眼跳幅度”“注视点个数”“AOI 访问次数”;处于中等效应值水平的依次是“注视时间占比”“瞳孔大小”“注视持续时间”“眨眼时间”“AOI 进出次数”“眨眼频率”“眨眼次数”。

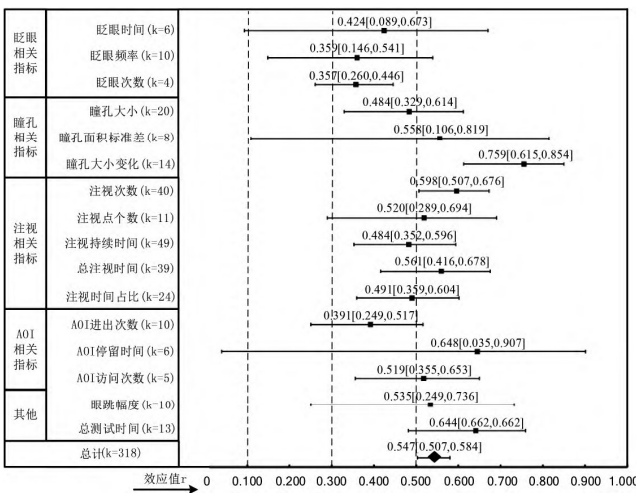


图 2 有效测量指标森林图

表 3 认知负荷来源对认知负荷测量指标的影响

认知负荷来源	数量	效应值	标准误	方差	95%置信区间		Z_r	p 值	组间效应值
					下限	上限			
任务难易程度	165	0.561	0.055	0.003	0.483	0.630	11.523	0.000	$Q=1.155$ $p=0.561$
学习材料设计方式	106	0.506	0.069	0.005	0.399	0.599	8.108	0.000	
学生学业水平	47	0.587	0.104	0.011	0.439	0.704	6.524	0.000	

森林图提供了可视化结果,能够快速观察得出眼动指标测量认知负荷的能力^[32]。将测量指标结果输出为森林图,显示效应值结果和对应的 95%置信区间,如图 2 所示。

(三)其他调节变量的效应分析

以任务的知识性质作为调节变量,计算出陈述性知识对应大的效应值 $r=0.523$ ($CI=[0.441,0.597]$),程序性知识对应大的效应值 $r=0.572$ ($CI=[0.492,0.642]$)。通过比较效应值发现,虽然程序性知识组的效应值大于陈述性知识组,但是两种类型的任务之间没有显著差异($Q=0.763, p=0.382>0.05$)。

以眼球追踪仪器的采样频率作为调节变量,计算出 60Hz 以下对应大的效应值 $r=0.733$ ($CI=[0.565,0.843]$),60Hz 及其以上对应大的效应值 $r=0.530$ ($CI=[0.470,0.586]$),异质性检验统计量 $Q=4.890$ ($p=0.027<0.05$),表明效应值在高低采样率之间存在显著性差异。同时,比较效应值发现,60Hz 以下的低采样率眼动仪的效应值大于 60Hz 及以上的高采样率眼动仪,说明低采样率的眼动仪对认知负荷的改变更加敏感,更有可能适用于多媒体学习领域的认知负荷测量。

以认知负荷来源作为调节变量,计算出任务难易程度对应大的效应值 $r=0.561$ ($CI=[0.483,0.630]$),学习材料设计方式对应大的效应值 $r=0.506$ ($CI=[0.399,0.599]$),学生学业水平对应大的效应值 $r=0.587$ ($CI=[0.439,0.704]$),三个调节变量都具有较大的效应值。比较效应值发现,学生学业水平组的效应值最高,其次是学习材料设计方式组,最低的是任务难易程度组,三者之间没有显示出显著差异($Q=1.155, p=0.561>0.05$),且都具有大效应值水平,说明眼动实验可以有效测量这三种不同来源的认知负荷眼动指标,具体内容见表 3。

五、研究结论

整体上,眼动仪各项指标受认知负荷的综合影响较为显著(总效应值 $r=0.547$),其测量认知负荷的有效性得到确认。通过分析得出如下结论:

(一)具有高效应量的眼动指标

本研究验证了眼动仪部分指标测量认知负荷的

有效性, 得出了若干项具有大效应值的眼动测量指标, 瞳孔、注视、眨眼、AOI 等相关指标更宜选取。

通过元分析发现, 瞳孔相关测量指标尤其是“瞳孔大小变化”对认知负荷具有较强的敏感性, 可能的原因是瞳孔的变化是不受实验对象主观控制的自主神经系统的活动。在高难度任务中, 由于任务的完成需要更多的认知资源和注意力, 大脑会自发地通过增加瞳孔的大小和扩张速度来调节视网膜上光线的入射量, 以提高视觉清晰度和视觉敏感度, 从而更好地应对任务要求。这种机制可以被解释为大脑在高难度任务中对视觉系统的适应性调节^[10]。相反, 在低难度任务中, 由于任务的完成需要的认知资源较少, 大脑对视网膜上光线的入射量的调节不需要那么明显, 因此, 瞳孔反应较弱^[33]。同时, 为了保证实验数据的准确性, 需要在实验过程中控制适度的环境光线并确保被试者在整个实验过程中保持相对稳定的头部姿势和注视距离。

通过研究发现, 注视相关测量指标尤其是“注视次数”对认知负荷具有较强的敏感性, 可能是因为注视过程涉及视觉和注意力两个认知过程, 而这两个认知过程都受到认知负荷的影响^[34]。通过记录实验对象注视过程的时间、频率、持续时间等指标, 可以反映其不同时间段内的注意力变化以及视觉加工的深度和广度、信息获取和处理的速度等, 这些信息可以帮助研究者更好地理解和分析认知负荷的复杂性, 提高研究的可靠性和有效性^[35]。因此, 这种视觉和注意力的双重效应使得注视相关测量指标能够提供更加精细和全面的认知负荷测量。

另外发现, “眨眼时间”“眨眼频率”“眨眼次数”受认知负荷影响较大。和瞳孔的变化相似, 眨眼也作为无意识的生理反应与注意力和认知负荷密切相关。在认知加工的过程中, 人们倾向于增加视觉固定的时间以减少干扰, 这会导致眨眼频率的降低, 同时, 随着任务时间的增长, 较低的眨眼频率导致眼睛不能保持湿润, 不能及时调节视觉焦距和减少眼睛的疲劳感, 其又会反作用于人们记忆和认知的表现^[36]。同样, 因为视觉固定的增加, 眨眼的的时间可能会被压缩以便更好地维持注意力和处理视觉信息, 因此, 眨眼时间在视觉需求增加时也会显著减少^[3]。

兴趣区的眼动指标尤其是“AOI 停留时间”也呈现出了较大的效应值, 说明通过划分不同兴趣区可以更充分地了解实验对象的认知负荷水平和认知加工过程的变化情况。例如: 当认知负荷增加时, 实验对象的注意力集中程度可能会下降, 从而导致其注视兴趣区的时间和次数减少, 这些眼动指标的变化可以反映

出不同任务难度下的认知负荷水平, 并提供有价值的信息。所以通过分析 AOI 兴趣区的相关眼动指标有助于研究者深入理解认知负荷对视觉加工的影响, 并进一步探究认知负荷的相关机制。

其余的测量指标虽然样本量充足, 但是经过分析被判定为不宜使用。例如: 返回眼跳(眼跳侵扰)可能受到多种因素的影响, 包括注意力的变化、情绪的波动、疲劳和认知策略的改变等。若研究必须使用这些指标, 为了提高这些指标测量结果的稳定性, 研究者可以采用重复测量、限制参与者的行为、控制环境和使用适当的数据处理方法等进行控制。

(二)调节变量的作用

陈述性知识性质和程序性知识性质两组之间虽然并未观察到效应值的显著不同, 但是两者之间存在一定的差异性。通过对比文献的研究细节发现, 当任务的知识性质是陈述性知识时, 学习者的眼动行为更倾向于关注阅读材料的表面特征, 容易出现回读、跳读等现象, 研究者也更愿意使用眼跳、眨眼等相关指标。相比之下, 当学习者学习程序性知识时, 他们的眼动行为更加关注步骤和顺序, 会更长时间地注视程序的步骤和细节, 研究者对此更倾向于使用注视、瞳孔等相关指标。这很有可能造成程序性知识和陈述性知识观察到的调节效应不明显, 未来需要更多研究的测量指标结果来提高效应值的统计功效和准确性。此外, 本研究提取到的使用陈述性知识作为学习材料的研究相较程序性知识大多缺乏反馈功能, 缺乏反馈可能会导致学生进行陈述性知识学习的认知负荷激增, 缩小了与有系统反馈的程序性知识组的差异, 从而影响实验结果^[37]。因此, 在进行多媒体学习环境及任务设计中, 要重视多媒体学习任务系统的反馈功能, 以防此因素影响实验结果。

在眼动仪采样率调节变量分析中, 60Hz 以下采样率与 60Hz 及以上采样率的眼动仪效应值水平有较大差异, 60Hz 以下采样率的眼动仪可能更适合用于测量多媒体学习领域的认知负荷。虽然理论上更高的采样率往往意味着眼动仪性能更好, 能产生更多的数据, 但是在实际使用中, 必须要根据自身需求选取一定采样率的眼动仪^[19]。在分析中发现, 使用高采样率眼动仪的研究者倾向于捕捉眼跳幅度等与复杂的脑神经科学相关的微眼跳指标^[38]。虽然此类测量指标潜力巨大, 目前的相关研究发现该指标处于大效应值水平, 但是高采样率组实验的效应值显著低于低采样率组实验的效应值。

在认知负荷来源调节变量分析中, “任务难易程

度”“学习材料设计方式”“学业水平”三者之间未呈现出显著的效应值差异,但是都具有大效应值水平。这表明在通过眼动实验测量认知负荷的实验中,不管研究者是通过改变实验组的任务难易程度,还是改变学习材料设计方式,抑或是根据学业水平划分实验对象,三种方式对不同组进行认知负荷的测量都是合理的。也就是说,在多媒体学习领域,眼动实验可以有效帮助研究者测量受不同因素影响的认知负荷。

六、结 语

本研究聚焦多媒体学习环境中认知负荷的眼动测量指标及其调节变量的研究,通过元分析方法提

取了若干具有较大效应量的眼动测量指标,并对知识类型、眼动仪的采样频率、认知负荷来源等调节变量进行了分析。相关研究结论能够为多媒体学习环境中认知负荷的有效测量提供依据与借鉴。虽然本研究中纳入的研究总数量较大,但并不是每一组眼动指标变量都有足够的样本支撑其效应值,这很可能是由于一些研究者只报告显示出明显统计学差异的测量指标,造成一些指标的有效性无法得到确认。因此,在后续的研究中需要设计认知负荷诱发试验以诱发不同层次的认知负荷,并以此为基础提取并构建能够有效表征学习者认知负荷的眼动指标,甚至是多模态测量指标。

[参考文献]

- [1] 王国华,聂胜欣,薛瑞鑫.多媒体学习中的认知负荷:测量方法与技术纵览[J].电化教育研究,2022,43(4):19-25,54.
- [2] WICKENS C D, HOLLANDS J G, BANBURY S, PARASURAMAN R. Engineering psychology and human performance [M]. 4th ed. Boston, MA: Pearson Education Inc., 2013:80-85.
- [3] WIERWILLE W W, EGGEMEIER F T. Recommendations for mental workload measurement in a test and evaluation environment[J]. Human factors, 1993, 35(2):263-281.
- [4] ZHANG L, ZHANG J. Eye-tracking technology in psychological research: progress and prospects [J]. Advances in psychological science, 2019, 27(4):655-666.
- [5] ROSENFELD M, JAHAN S, NUNEZ K, et al. Cognitive demand, digital screens and blink rate [J]. Computers in human behavior, 2015, 51:403-406.
- [6] BENEDETTO S, PEDROTTI M, MININ L, et al. Driver workload and eye blink duration [J]. Transportation research part f: traffic psychology and behaviour, 2011,14(3):199-208.
- [7] TUYVEN E, VAN DER GOTEN K, VANDIERENDONCK A, et al. The effect of cognitive load on saccadic eye movements [J]. Acta psychologica, 2000,104(1):69-85.
- [8] SALVUCCI D D, GOLDBERG J H. Identifying fixations and saccades in eye-tracking protocols [C]//Proceedings of the 2000 Symposium on Eye Tracking Research & Applications. New York: ACM, 2000:71-78.
- [9] KANG O E, HUFFER K E, WHEATLEY T P. Pupil dilation dynamics track attention to high-level information [J]. PLOS One, 2014,9(8):1-6.
- [10] KRZYSZTOF K A, ANDREW T D, ANNA N, et al. Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze[J]. PLOS One, 2018,13(9):1-23.
- [11] MARQUART G, CABRALL C, DE WINTER J. Review of eye-related measures of drivers' mental workload [J]. Procedia manufacturing, 2015,3:2854-2861.
- [12] SAVAGE S W, POTTER D D, TATLER B W. Does preoccupation impair hazard perception? A simultaneous EEG and eye tracking study[J]. Transportation research part f: traffic psychology and behaviour, 2013,17(2):52-62.
- [13] ZAGERMANN J, PFEIL U, REITERER H. Measuring cognitive load using eye tracking technology in visual computing[C]//Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization. New York: IEEE, 2016:78-85.
- [14] GRANHOLM E, VERNEY S P. Pupillary responses and attentional allocation problems on the backward masking task in schizophrenia[J]. International journal of psychophysiology, 2004,52(1):37-51.
- [15] JOHNSON C I, MAYER R E. An eye movement analysis of the spatial contiguity effect in multimedia learning [J]. Journal of experimental psychology: applied,2012,18(2):178-191.

- [16] 辛自强,林崇德.认知负荷与认知技能和图式获得的关系及其教学意义[J]. 华东师范大学学报(教育科学版),2002,20(4):55-77.
- [17] NECHES R, LANGLEY P, KLAHR D. Learning, development, and production systems [M]// NECHES R, LANGLEY P, KLAHR D. Production system models of learning and development. Cambridge: The MIT Press. 1987:1-45.
- [18] WIERTS R, JANSSEN M J A, KINGMA H. Measuring saccade peak velocity using a low-frequency sampling rate of 50 Hz[J]. IEEE transactions on biomedical engineering, 2008,55(12):2840-2842.
- [19] ANDERSSON R, NYSTRÖM M, HOLMQVIST K. Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more[J]. Journal of eye movement research, 2010,3(3):1-12.
- [20] 李东.先验知识、归纳水平、逻辑动画和材料性质对录屏式微课学习的认知负荷与学习成绩的影响[D].福州:福建师范大学,2017.
- [21] 胡晓玲,袁民,赵凌霄.2002—2022 年间我国学习主题元分析研究的质量评价[J].电化教育研究,2023,44(1):106-115.
- [22] ROSENTHAL R, DIMATTEO M R. Meta-analysis: recent developments in quantitative methods for literature reviews [J]. Annual review of psychology,2001,52(1):59-82.
- [23] COBAN M, BOLAT Y I, GOKSU I. The potential of immersive virtual reality to enhance learning: a meta-analysis [J]. Educational research review, 2022,36(3):1-19.
- [24] 叶爱敏.整合微视频的数字化阅读眼动行为研究[D].温州:温州大学,2016.
- [25] ROP G, VAN WERMESKERKEN M, DE NOOIJER J A, et al. Task experience as a boundary condition for the negative effects of irrelevant information on learning[J]. Educational psychology review,2018,30(10):229-253.
- [26] TOKUDA S, OBINATA G, PALMER E, et al. Estimation of mental workload using saccadic eye movements in a free-viewing task[C]// 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. New York: IEEE, 2011: 4523-4529.
- [27] YANG J, WANG S, TONG X, et al. Semantic and plausibility effects on preview benefit during eye fixations in Chinese reading[J]. Reading and writing,2012,25(9):1031-1052.
- [28] ONNASCH L, WICKENS C D, LI H, et al. Human performance consequences of stages and levels of automation: an integrated meta-analysis[J]. Human factors,2014,56(3):476-488.
- [29] HIGGINS J P T, THOMPSON S G, DEEKS J J, et al. Measuring inconsistency in meta-analyses[J]. BMJ, 2003,327(7414):557-560.
- [30] 卫旭华.组织与管理研究中的元分析方法[M].北京:科学出版社;2021:48.
- [31] COHEN J. A power primer[J]. Psychological bulletin, 1992,112(1):155-159.
- [32] SCHAEFER K E, CHEN J Y C, SZALMA J L, et al. A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems[J]. Human factors, 2016,58(3):377-400.
- [33] HESS E H, POLT J M. Pupil size as related to interest value of visual stimuli[J]. Science,1960,132(3423):349-350.
- [34] VANNESTE P, RAES A, MORTON J, et al. Towards measuring cognitive load through multimodal physiological data[J]. Cognition, technology & work, 2021,23(3):567-585.
- [35] JACOB R J K, KARN K S. Eye tracking in human-computer interaction and usability research: ready to deliver the promises[M]// RADACH R, HYONA J, DEUBEL H. The mind's eye: cognitive and applied aspects of eye movement research. Amsterdam: Elsevier, 2003,2(3):573-605.
- [36] HAYES J, ALLINSON C W. Cognitive style and its relevance for management practice [J]. British journal of management, 1994,5(1):53-71.
- [37] CLARK R E. Media will never influence learning[J]. Educational technology research and development, 1994,42(2):21-29.
- [38] EBERHARD K M, SPIVEY-KNOWLTON M J, SEDIVY J C, et al. Eye movements as a window into real-time spoken language comprehension in natural contexts[J]. Journal of psycholinguistic research, 1995,24(6):409-436.

Research on Eye-tracking Indicators for Cognitive Load Representation in Multimedia Learning

WANG Guohua¹, TIAN Lianghao¹, NIE Shengxin¹, ZHU Ke², LIANG Yunzhen¹

(1.Department of Education, Henan Normal University, Xinxiang Henan 453000;

2.Intelligent Education Henan Collaborative Innovation Center, Xinxiang Henan 453000)

[Abstract] Measuring cognitive load through eye-tracking devices has become an important method for cognitive load measurement. Although many studies have relied on the measurement data output by eye-tracking devices to characterize cognitive load, there are no clear indicators of which eye-tracking data can accurately characterize the cognitive load status of learners. This study used a meta-analysis method of evidence-based research to quantitatively analyze 62 experimental studies at home and abroad, to systematically examine the impact of cognitive load on eye-tracking indicators and explore eye-tracking indicators that can characterize cognitive load in multimedia learning. It is found that: (1) the overall effect size of the samples included in the study is 0.547 ($k=318$), indicating that eye-tracking indicators are significantly affected by cognitive load; (2) analyzing the different measurement indicators as moderating variables, it is found that 15 eye-tracking indicators with significant effect sizes, such as pupil diameter, can better characterize learners' cognitive load status; (3) in the analysis of other moderating variable, it is found that the eye-tracking device with a sampling rate below 60Hz is more suitable for the measurement of cognitive load in multimedia learning, while there is no moderating effect for the nature of knowledge and the source of cognitive load.

[Keywords] Cognitive Load; Multimedia Learning; Eye-tracking Experiment; Eye-tracking Technology; Meta-analysis

欢迎订阅 2024 年《电化教育研究》

《电化教育研究》是中文核心期刊、CSSCI 来源期刊、RCCSE 中国权威学术期刊、AMI 核心期刊、复印报刊资料重要转载期刊,素有“中国电化教育理论研究基地”之称誉。主要研究现代教育前沿问题,服务全国教育教学改革;关注国内外信息化教育理论的创新与发展。

《电化教育研究》开设的栏目主要包括理论探讨、教育数字化、网络教育、学习环境与资源、课程与教学、学科建设与教师发展、历史与国际比较、中小学电教等。

聚焦学术前沿、追求学术品位、体现学科特色、支持学术争鸣和学术创新。

《电化教育研究》为月刊(每月 1 日出版),全彩印刷,国际标准大 16 开本,128 页,单价 15 元/期,全年 180 元。欢迎广大读者到当地邮局订阅或扫码在线订阅。邮发代号:54-82。

联系电话:0931-7971823,0931-7970887

官方网站: <http://aver.nwnu.edu.cn>

微信公众号: e-EducationResearch(电化教育研究)



(请扫码订阅)