

多媒体学习中的认知负荷:测量方法与技术纵览

王国华, 聂胜欣, 薛瑞鑫

(河南师范大学 教育学部, 河南 新乡 453000)

[摘 要] 认知负荷是在线学习及多媒体学习领域研究的焦点问题,如何科学、有效地测量学习者的认知负荷是相关研究开展的基础,更是减少学习者认知资源消耗、提升学习成效的关键。文章以国内外近 20 年来的 301 篇实证研究论文为基础,使用客观—因果二维四象限分类法对不同类型的认知负荷测量方法与技术进行分类,并对相关测量方法与技术的理论假设、优点及不足进行分类评述,并得出些许启示。通过分析发现,不同测量方法各有优劣并具有互补性,基于认知科学和信息科学的测量方法值得重点关注,深度学习技术会成为未来认知负荷测量的关键技术。

[关键词] 多媒体学习; 在线学习; 认知负荷; 测量方法; 测量技术

[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民族教育信息化研究。E-mail: wgh19892008@126.com。

一、引 言

在线学习具有场域开放、复杂度高、信息冗余的特点,媒体及资源的不合理使用可能引起学习者的混乱和认知超载,影响学习者学业表现。认知负荷是学习者在进行信息处理与加工过程中所消耗的认知资源^[1]。认知负荷作为多媒体学习中最为重要的认知要素,对多媒体学习效果产生重要影响,如何准确衡量认知负荷是当前研究迫切需要解决的关键问题。

二、认知负荷与多媒体学习

人类在学习和问题解决过程中不可避免地要消耗认知资源,而个人单位时间内所能处理和加工的信息元素是有限的,当处理信息所需的认知资源数量超过个人工作记忆资源容量时就会出现认知超载现象。

多媒体环境下的学习涉及不同类型技术媒介的使用,不同的教学媒介、资源呈现方式及各类线索的应用使得学习者的认知负荷难以得到有效控制,因

此,多媒体学习环境下的认知负荷问题更值得研究。Mayer 对多媒体学习领域中学习者的认知过程进行了探索,并总结出著名的多媒体学习理论以均衡学习者的认知负荷水平。多媒体学习理论的基础是双编码理论假设和双通道理论假设。多媒体学习的生成理论将这两个假设与生成学习方法相结合,指出学习者从学习材料中主动选择相关的视觉和语言信息,并通过在视觉和语言工作记忆之间建立联系,分别将其存储在视觉和语言工作记忆中,通过分配认知资源和工作记忆容量建立参照链接,整合心理表征和先验知识,当学习者接收到的信息数量超出工作记忆容量时就会产生认知负荷效应^[2]。

当前,多媒体学习领域使用的认知负荷测量方法包括主观测量法、行为分析法、双任务分析法、绩效分析法、眼动追踪法以及生理测量法等。多种方法的应用为认知负荷测量的客观性、准确性提供了可能。本研究通过文献分析的方法,梳理多媒体学习领域中使用的认知负荷测量方法及相关指标,分析各类方法的

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);2020 年度河南省哲学社会科学研究青年项目“在线学习社区中知识共享的动力机制及仿真研究”(项目编号:2020CJY042)

优点及不足,以期为后续研究提供借鉴。

三、研究设计

(一)数据来源

本研究所选用的数据来自 Web of Science、Education Resources Information Center (ERIC)、EBSCO 以及国内的 CNKI 等数据库。将数据的时间检索跨度设定为 2001 年 1 月 1 日至 2021 年 9 月 10 日。使用关键词“多媒体”+“认知负荷”、“多媒体学习”+“认知负荷”、“Multimedia learning”+“Cognitive load”、“Multimedia”+“Cognitive load”进行检索。文献来源筛选标准为:SSCI 索引、SCI 索引、SCIE 索引、EI 索引、CSSCI 索引。内容筛选标准为:涉及多媒体学习中的认知负荷测量、正式刊发的量化学术论文或会议论文。通过筛选得到有效文献 301 篇。

(二)分类框架

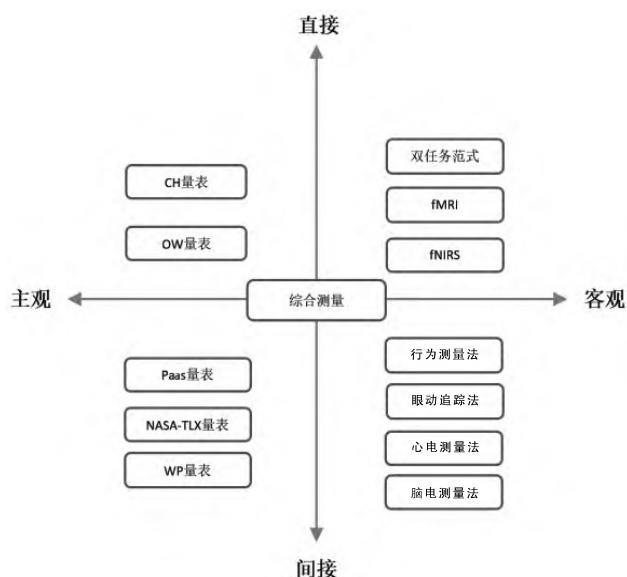


图1 认知负荷测量方法分类框架

当前测量认知负荷的方法和工具众多,因此,我们需要建立一个分析框架对认知负荷测量方法进行类属的划分,以便研究者能更加清晰地了解不同方法的特征。目前常用的认知负荷评估方法可以从两个维度进行划分,即主客观性和因果关系。从测量方法的主客观性上来看,可以分为主观测量法和客观测量法;从测量方法的因果关系上来看,可以分为直接测量法和间接测量法。上述两种分类方法都属于二分类法,虽然分类形式简单明了,但却无法从其他维度更进一步分析每种方法的优劣。而将两种方法结合起来,使用主客观性和因果关系两个维度建立四象限分类方法,有助于我们更加清楚地对认知负荷测量方法进行描述。本研究将在文献梳理的基础上,对各

项文献中所出现的认知负荷测量方法及工具进行归类划分,分类框架如图1所示。

(三)分析指标

在对多媒体学习环境中认知负荷测量方法及内容进行分类、总结与梳理的基础上,从哪些方面分析认知负荷测量的有效性和适切性,是需要关注的重要问题。Eggemeier 提出了判断认知负荷测量方法恰当性和有效性的准则。具体而言,包括以下七个指标:灵敏度、诊断性、干扰性、效度、信度、便利性、测量对象接受度^[3]。

四、研究结果

(一)直接—主观类测量方法

主观测量是认知负荷研究中最常用的评估方法,一般使用“你认为学习材料的难度如何?”的方式进行提问与测量。主观测量法的应用基于这样的假设,即个人可以评估自己的认知过程,对自身在完成任务过程中所消耗的认知资源进行自主评级^[4]。在直接—主观类测量方法中,使用单一维度量表让学习者陈述学习材料困难度或负荷水平的可接受程度。常用的测量工具包括 CH 量表(Cooper-Harper Scale)和 OW 量表(Overall Workload Scale)。CH 量表自 1969 年开发之后,一直用于测量人们的脑力负荷,研究者将其迁移到认知负荷测量中,发现该量表也具有较高的信度及效度。CH 量表将学习者操作性水平等级分为 10 级,并具有对应的定义解释,等级越高所带来的认知负荷量越高^[5]。OW 量表在使用过程中,让使用者根据学习过程中的整体负荷量进行主观评分,分值越高表明认知负荷量越高^[6]。

直接—主观类测量方法使用单一维度的测量量表进行学习者认知负荷测量,这种方法具有较高的敏感性且易于使用,在使用过程中不具有侵入性,容易被测量对象所接受^[7]。但是,由于这种评分方法主观性过强,并且没有解释性因素阐释认知负荷水平变化的原因,所以这种测量方法的弊端也是显而易见的。学习者认知负荷水平差异也可能是由任务难度、学习者的个人能力水平或不同的注意过程引起的。另外,自主评分通常是在学习活动之后进行的,因此,主观测试等级量表不能提供学习过程中认知负荷变化的连续信息。一维量表只能用来测量学习者认知负荷的水平,而无法对相关数据作出解释,适用场景较窄,因此,在多媒体学习研究中,这种方法应用较少。

(二)直接—客观类测量方法

直接—客观类测量方法兼具认知负荷测量的直

接性和客观性,是目前测量认知负荷最为有效和精确的方法类型。其中一种直接—客观测量认知负荷的方法是实验心理学中经常出现的双任务范式。双任务范式基于有限认知资源的假设,而这种资源可以灵活地分配给任务解决的不同方面^[8]。双任务范式是工作记忆研究中使用的重要方法,实验过程中被试对象需要在进行主任务活动的同时(如观看视频或阅读文本等),完成一项次任务(如分辨音频、打拍子、分辨图片等)。次任务的效果与空闲的工作记忆容量直接相关,空闲容量本身取决于主要任务所需的容量^[9]。所以,为了使次任务成为一个具有敏感性的衡量指标,它需要与主要任务分配相同的认知资源,否则次任务与主任务将不具有关联性。此外,次任务必须要简单可行,以免抑制同步学习过程。正如认知负荷理论所认为的,如果特定学习材料引起的额外负荷量取决于该材料的教学设计,那么同一内容的不同设计变量会导致次要任务的不同表现^[10]。但是,双任务实验范式需要研究者具有扎实的实验心理学基础,对实验程序及试验场地要求较高。并且,在研究过程中,次任务会施加额外的认知负荷干扰主要任务,双任务范式由于较高的学习活动侵入性受到批评。

使用脑成像技术观测任务执行过程中大脑影像的变化是具有广阔前景的认知负荷测量方法。脑成像技术测量认知负荷的理论假设是“大脑不同区域负责不同的工作记忆功能,学习任务执行过程中额外的认知负荷会导致大脑的某些区域发生改变”^[11]。多媒体学习中使用的脑成像技术一般包括功能性磁共振成像(fMRI)和功能性近红外光谱技术(fNIRS)。功能性磁共振成像(fMRI)通过对不同认知状态下大脑活动区域的活跃程度及变化的观测,建立参数化程序与认知负荷之间的关联性。O'Hare等人对高、中、低3个层次认知负荷进行了研究,发现额叶、顶叶和小脑区域具有显著性的变化^[12]。Cabeza等人的研究发现,前额皮质区域的布罗德曼(Brodmann)6区、9区、44区和46区也会随着执行学习任务的差异而变化^[13]。fMRI在识别认知负荷的准确率方面具有无与伦比的优势,但是其昂贵的价格及较大的体积在使用中带来了诸多不便,使得fMRI无法得到广泛应用。功能性近红外光谱技术(fNIRS)是一种非侵入性神经影像技术,它基于近红外光谱(700纳米~900纳米)中吸收光的变化来测量大脑的血液动力学反应^[14]。fNIRS作为一种脑活动成像技术,具有实时性特征,可以在学生学习过程中不间断收集其大脑血液动力学数据,实现认知负荷的全域检测。例如,Yazidi等人基于N-back实验范式成

功地诱发了学习者的不同层次的认知负荷状态,通过fNIRS技术对学习者的脑部血氧状态的测量,发现使用额叶皮层中的六个通道获得了86.20%的识别率,位于额叶区域的AFpz通道获得了86.34%的识别率^[15]。尽管fMRI研究在临床上被认为可以准确识别大脑生理活动,但设备的固有成本、缺乏便携性和低时效分辨率限制了可实验对象的数量和类型,相比之下,fNIRS提供了一种可靠的替代方法来收集学习者血液动力学数据,其时间分辨率可以达到fMRI的10倍。fNIRS技术成为认知负荷测量中颇具潜力的测量技术。

直接—客观类测量方法在有效识别认知负荷过程中体现出了无与伦比的识别灵敏度,并且这种认知负荷识别发生在学习者执行学习任务的过程中,避免了主观测量方法自我报告的滞后性,使得认知负荷的识别更加精确。但是,直接—客观类测量方法最大的问题在于其测量过程的侵入性过于明显,无论是脑成像技术还是双任务范式都会对学习者的学习任务产生影响。因此,这种测量方法只能适用于专业的心理研究实验室或脑科学研究实验室中的场景,认知科学及脑科学研究者对这类方法情有独钟,测量方法的易用性、可接受性等都不具有优势,适用范围较窄。

(三)间接—主观类测量方法

间接—主观类测量方法是属于主观测量法的一种,其假设前提是学习者的认知资源消耗与任务难度、心理努力、时间压力等要素具有紧密关联性^[16]。与直接—主观类测量方法中使用的单一维度量表不同,间接—主观类测量方法使用多维量表测量学习者的认知负荷水平,除了使用任务难度评级直接测评之外,还通过心理努力、投入、时间压力等要素从侧面衡量学习者认知负荷水平。

Paas自我评定多维量表是多媒体学习中认知负荷测量使用频率最高的工具,该量表包括任务难度和心理努力两个方面,任务难度直接表示学习者认知负荷量,心理努力则从侧面间接表示学习者认知负荷量^[17]。Paas量表因其简单易操作且具有较高的信效度而成为最受欢迎的认知负荷多维测量量表。

NASA-TLX(National Aeronautics and Space Administration-Task Load Index)自我评定量表是由美国国家宇航局开发用于测量宇航员脑力负荷量的多维量表工具,该量表包括绩效水平、努力程度、时间需求、心理需求、体力需求、受挫程度6个维度,每个维度分为20个等级,依据权重系数可计算出学习者认知负荷水平^[3]。NASA-TLX量表因其简便易用、测量的敏感性以及结果的可解释性受到多媒体学习研究者

的青睐,应用范围较广。

WP(the Workload Profile Index Ratings)自我评定量表以负荷多重资源模型为基础,涉及中枢处理、响应、空间编码、语言、视觉、听觉、操作输出、语言输出8种认知资源,完成学习任务后学习者分别对上述8种认知资源占用量进行赋值评分^[18]。WP量表使用方法简单,没有使用场景的限制,并且在认知负荷测量过程中具有较好的抗干扰性及敏感性,在多媒体学习研究中使用较为广泛。

间接—主观类测量方法在使用过程中具有简便性、易用性、非侵入性、易于被接受等优势。相较于直接—主观类测量方法而言,间接—主观类测量方法所使用的测量工具包含了更多的内容信息,除了可以对认知负荷进行直接测量,相关要素也可以对认知负荷变化的原因作出解释。另外,间接—主观类测量方法可以进行差异化任务类型之间认知负荷测量结果的对比,并且对不同等级认知负荷的变化比较敏感。针对特定维度的等级量表可以显著区分不同类型的认知负荷,这也是其他类型测量方法所不具备的优势。但是,此类方法最大的问题在于其主观性太强,评级尺度因客观程度、有效性和可靠性的问题而受到批评^[19]。学习者在自评过程中,不可避免地会出现所有自陈式量表的通病——“社会赞许效应”,进而影响结论的可靠性。虽然,间接—主观类测量方法中使用的自陈式量表具有一定的局限性,但是这种方法仍然是当前使用最为广泛、较为可靠的认知负荷测量方法,几乎适用于所有研究场景,具有极强的适用范围。

(四)间接—客观类测量方法

行为测量法是指通过分析学习者行为特征而测量其认知负荷的方法,行为测量法的前提假设是“不同学习任务或学习材料与个人行为具有关联性,随着任务难度等级的发展,学习者特征发生显著变化”^[20]。在线学习环境中一般使用鼠标轨迹及点击行为来衡量学习者认知负荷水平。例如,有些研究者提出了一种基于鼠标动力学技术的测量方法,通过研究发现学习者在高认知负荷下表现出更长的任务持续时间、更长的鼠标移动、更多的方向变化和较慢点击的速度^[21]。数字化环境下的书写特征也可以作为测量学习者认知负荷的要素,Luria等人使用数字化手写板研究了学习者在不同认知负荷水平中进行书写的特征,发现书写的时间(书写速度)、空隙(笔在空气中停留时间)和角速度可以作为测量认知负荷的有效指标^[22]。认知行为测量通过隐性的行为记录工具及方法,收集相关数据信息进而实现对学习者认知负荷的测量,具有较

强的客观性、隐蔽性、可接受性以及低侵入性。但是,行为分析法因存在较为复杂的技术难度和单一行为分析准确度较低等问题而未被广泛使用。

眼睛注视反映了注意力的过程,眼动追踪技术为多媒体学习环境下认知状态研究提供了有效方法^[23]。眼动追踪技术应用于认知负荷测量的前提假设是“不同难度等级的学习材料或任务必然使学习者的眼动轨迹及凝视时间发生改变”^[24]。多媒体学习领域一般通过注视、扫视来测量认知状态。Liu等人认为,可以从时间、空间和计数刻度3个尺度指标来衡量学习者的认知状态^[25]。除了上述常用指标之外,Krzysztof等人通过研究发现,瞳孔直径和微眼跳幅度的试验间变化似乎可以充分区分不同层次的认知负荷^[26]。Kontogiorgos发现,学习者的瞳孔直径会随着任务类型的难度改变而变化^[27]。眼动追踪法以其易用性、敏感性及客观性受到众多研究者的青睐,成为生理测量中应用最为广泛的技术类型,使用眼动追踪法来测量多媒体学习者的认知负荷得到了较为广泛的应用。

心电测量法测量学习者认知负荷的理论假设是心脏状态会随着人认知状态的变化而发生改变,使用心电信号可以表征这种变化。心电测量一般使用心脏周期、心率及心率变异性作为衡量认知负荷变化的尺度。从具体的分析指标而言,每个尺度可分为时域指标与频域指标。能够显著预测认知状态变化的时域指标包括心率间隔的间期均值、间期标准差、间期差值均方、间期大于50ms的个数等,相关的频域指标包括低频功率(0.04~0.15Hz)、高频功率(0.15~0.4Hz)、标准化的高频功率等^[28]。心脏测量具有其他生理测量法不具备的优势,例如:用于收集数据的设备相对便宜;设备操作相对比较简单,无须进行培训;侵入性较低;能够测量全过程数据。

使用脑电技术评估学习者认知负荷是近些年兴起的一种新型测量方法。相较于可以直接观测脑部活动的脑成像技术而言,脑电测量法是通过脑电波或认知电位来间接衡量学习者认知负荷的方法。脑电技术测量认知负荷的假设前提是不同的任务等级所对应的脑部神经点位会发生变化,这种变化是可记录、可理解的。具体而言,脑电技术可以分为脑电图(EEG)和事件相关电位(ERP)两种。Castro发现,在高工作记忆负荷条件下,使用 θ 波(4~7Hz)能够显著预测认知负荷的发生及强度^[29]。Ghani等人发现,随着认知困难程度的增加,ERP各成分的振幅也发生变化,主要指标包括N100、N200、P200、P300等不同ERP成分^[30]。脑电测量法以其在认知状态评估中的可靠性、敏感性等

特征受到多媒体学习研究领域的认可,成为认知测量的重要方法。

间接—客观类测量方法具有其他测量方法所不具有的优势,体现在以下几点:在认知负荷测量过程中具有颇高的稳定性;对认知负荷水平的变化具有较高的敏感性;测量过程具有较低的侵入性,不会影响学习者的主任务;抗干扰性更强,不会因为学习者个人内部特征而影响测量的可信度。但是,这种方法的劣势也是显而易见的:测量方法、程序以及技术工具需要研究者具有较高的操作基础;对测量的环境要求较高,必须在特定实验室完成;被实验对象接受度稍低。从适用场景来看,小型化的仪器设备推动科学研究逐渐走出实验室进入真实课堂,使得这种方法类型适用于多种研究场景。而从研究群体来看,认知科学、信息科学及教育科学研究者都在关注间接—客观类测量方法的使用,该类型方法应用广泛度会越来越高。

(五)综合测量方法

综合测量方法与其他测量方法有所不同,该方法并没有明确的技术路线与评测指标,是属于叠加性的测量方式,即将不同测量技术叠加使用以实现认知状态评估的多维校准,从而提升认知负荷测量的可靠性和准确性。事实上,使用综合测量方法来测量多媒体学习中的认知负荷现象已经较为普遍,并且综合测量方法对认知负荷测量的精确度已经超越单一测量技术的精确度,例如在认知负荷测量中同时使用两种或两种以上的主观量表能得到更加准确的结果。李金波在对人机交互过程中的认知负荷进行研究中发现,将Paas量表与眼动追踪技术相结合,使用机器学习算法对认知负荷的识别率达到78%^[31]。Wang等人通过对心率变异性(HRV)和脉搏率变异性(PRV)数据特征的提取与融合,发现这种方法能够实现高负荷状态的准确识别^[32]。

综合测量法的优势在于综合了不同测量技术的特点,可以实现认知负荷测量的可靠性与准确性。但是这种方法的弊端也是显而易见的,操作程序与技术要求过高,尤其是涉及不同模态数据的测量需要进行多模态数据的融合与计算,这就涉及深度学习算法方面的知识,教育领域的研究者在面对上述问题时显得比较被动。因此,虽然该方法在认知负荷的识别精度及适用范围方面都具有天然的优势,但是,数据科学的高门槛严重阻碍了研究者对该方法的使用。

五、研究结论及启示

(一)研究结论

通过文献梳理,本研究总结与分析了各类型认知负荷测量方法及技术工具的优缺点,表1为不同方法横向比较的结果。

(二)研究启示

1. 不同测量方法各有优劣,具有互补性

尽管学习者自主报告的测量形式因为容易受到主观因素的影响而受到批评,但是主观测量法依然是认知负荷测量的基础方法,在所有方法类型中占据最大比重。得益于主观测量法使用过程中较高的便利性、可接受度以及较低的使用门槛,尤其多维量表在认知负荷诊断性方面的天然优势,主观测量法仍然是众多研究者的首选方法。Merat的研究表明,在严格控制变量的环境中,主观评级法与生理测量法在识别不同层次难度任务中具有同等的敏感度^[33]。并且,基于认知及数据科学的测量技术与方法严重依赖各种行为及生理信息采集设备,并且在应用过程中存在较强的侵入性,应用与推广也受到如生理测量设备使用具有较高的技术要求等诸多因素制约,使用人员需要接受充分的培训。此外,诸如fMRI等过于昂贵的设备,使其难以被推广应用。因此,主观测量法较高的性价比使其无法被替

表1 认知负荷测量方法与技术比较

方法类型	具体方法	灵敏度	诊断性	干扰性	效度	信度	便利性	接受度
直接—主观类测量方法	一维量表	中	低	低	中	中	高	高
直接—客观类测量方法	双任务范式	高	低	高	高	高	中	高
	fMRI	高	中	高	高	高	低	低
	fNRIs	高	中	中	高	高	中	中
间接—主观类测量方法	多维量表	中	高	低	中	中	高	高
直接—客观类测量方法	行为测量法	中	低	低	高	高	高	高
	眼动追踪法	高	低	中	高	高	中	中
	心电测量法	中	低	中	高	高	中	中
	脑电测量法	高	中	中	高	高	中	中
综合测量方法		高	高	中	高	高	中	中

代,各类测量方法多元共存的局面会一直存在。

2. 基于认知科学和信息科学的测量方法值得重点关注

虽然认知行为和脑科学的研究严重依赖标准化的试验场景,但是该类型方法在揭示学习者认知机制及状态方面具有无与伦比的优势。从文献数量及发文时间来看,近年来呈现出迅猛的增长态势。不同于主观测量法,生理测量以客观生理数据作为衡量认知负荷层级的尺度,生理测量的客观性特征使其可以避免主观经验造成的偏差。此外,生理测量方法的另一优势是测量的过程性及连续性,可以不间断收集与分析学习者的生理数据,实现认知负荷的全域测量,更加客观准确。

从测量方法的应用状况来看,行为测量法并未受到研究者的青睐,这可能是由于行为测量法在认知负荷识别中的灵敏度不足造成的。当前的研究更多地关注实验环境下多媒体学习认知负荷的测量工作,而对开放性环境下的研究略显不足,这可能与认知实验需要进行严格的变量控制有关。然而,常规课堂环境以及大规模在线学习环境中的认知负荷测量对于教学活动的开展更具有实践意义,并且在开放性环境中不可能使用认知实验的形式对学习者的认知负荷进行测量。在此种情形下,通过学习者行为数据分析继而衡量学习者认知状况的方法将是值得关注的研究方向。具体而言,基于行为数据分析学习者认知负荷可分为课堂教学和在线学习两个方向。课堂教学中的认知负荷识别是基于计算机视觉与数据技术实现的行为识别,在预先建立学习者行为与认知负荷关联模型

的基础上,通过深度学习技术实现学习者认知状态的诊断。大规模在线学习环境中的认知负荷识别,需要在基于海量数据分析基础上构建在线学习行为与认知负荷状态的关联模型,使用深度学习算法(如注意力网络)评估学习者的认知状态。开放场景下基于行为数据分析的认知负荷测量具有必要性和紧迫性,显然,上述两种分析方法需要有较强的信息科学与认知科学研究背景,当前认知负荷的测量以认知科学理论为基础,相关研究者主要以认知科学背景为主,单一学科背景不利于认知负荷测量方法及技术的发展。

3. 深度学习技术会成为未来认知负荷测量的关键技术

综合测量方法结合了诸多测量方法的优势,兼具主观性与客观性、直接性与间接性等方法的优点,使得认知负荷的测量工作更为精确可靠。从文献研究来看,综合测量的内容已经从原来的“量表结合”向“跨模态结合”发展,计算方法也从原来的“简单叠加”向“多模态数据融合”发展,可以使研究的可靠性、精确性更上一层楼。无论是基于行为数据的认知负荷识别与分析,还是跨模态数据的分析,都需要进行大量的数据融合与计算,其背后的核心技术都是深度学习技术。在研究过程中可能会涉及图像模态表情与行为识别技术、多模态数据对齐与融合技术、多模态数据分类识别技术等,具体到算法可能会涉及随机森林、支持向量机、深度神经网络等分类模型等。未来认知负荷测量方法纵深发展的核心在于数据,而数据分析与计算的关键在于深度学习技术,因此,深度学习技术会成为未来认知负荷测量的关键技术。

[参考文献]

- [1] 赵立影.基于工作记忆的多媒体学习设计[J].电化教育研究,2011(8):98-102.
- [2] 李晶,郁舒兰,金冬.均衡认知负荷的教学设计及知识呈现[J].电化教育研究,2018,39(3):23-28.
- [3] 孙崇勇,刘电芝.认知负荷主观评价量表比较[J].心理科学,2013,36(1):194-201.
- [4] ANMARKRUD F, ANDRESEN A, BRATEN I. Cognitive load and working memory in multimedia learning: conceptual and measurement issues[J]. Educational psychologist, 2019, 54(2): 1-23.
- [5] COOPER G. Cognitive load theory as an aid for instructional design[J]. Australasian journal of educational technology, 1990, 6(2):108-113.
- [6] HILL S G, IAVECCHIA H P, BYERS J C, et al. Comparison of four subjective workload rating scales[J]. Human factors: the journal of the human factors and ergonomics society, 1992, 34(4):429-439.
- [7] WANG J, ANTONENKO P, KEIL A, DAWSON K. Converging subjective and psychophysiological measures of cognitive load to study the effects of instructor-present video[J]. Mind brain and education, 2020, 14(3): 279-291.
- [8] WIRZBERGER M, HERMS R, BIJARSARI S E, et al. Schema-related cognitive load influences performance, speech, and physiology in a dual-task setting: a continuous multi-measure approach[J]. Cognitive research: principles and implications, 2018, 46(3):1-16.
- [9] MEHLER B, REIMER B, COUGHLIN J F. Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: an on-road study across three age groups[J]. Human factors, 2012, 54(3):396-412.

- [10] KORBACH A, BRÜNKEN R, PARK B. Measurement of cognitive load in multimedia learning: a comparison of different objective measures[J]. *Instructional science*, 2017, 45(8):515-536.
- [11] 理查德·克拉克,文森特·克拉克,姜敏.从教育新行为主义到神经科学:以认知负荷研究的起源与未来为视角[J].*现代远程教育研究*,2014(3):52-65.
- [12] O'HARE E D, LU L H, HOUSTON S M, BOOKHEIMER S Y, SOWELL E R. Neurodevelopmental changes in verbal working memory load-dependency: an fMRI investigation[J]. *NeuroImage*, 2008, 42(4):1678-1685.
- [13] CABEZA R, NYBERG L. Imaging cognition II: an empirical review of 275 PET and fMRI studies [J]. *Journal of cognitive neuroscience*, 2000, 12(1):1-47.
- [14] KIM H Y, SEO K, JEON H J, et al. Application of functional near-infrared spectroscopy to the study of brain function in humans and animal models[J]. *Molecules & cells*, 2017, 40(8):523-532.
- [15] YAZIDI A, PINTO-ORELLANA M A, HAMMER H, et al. Solving sensor identification problem without knowledge of the ground truth using replicator dynamics[J]. *IEEE transactions on cybernetics*, 2020, 52(1):16-24.
- [16] LEPPINK J, PAAS F, VLEUTEN C P M V D, et al. Development of an instrument for measuring different types of cognitive load[J]. *Behavior research methods*, 2013, 45(4):1058-1072.
- [17] PAAS F G. Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach [J]. *Journal of educational psychology*, 1992, 84(4):429-434.
- [18] TSANG P S, VELAZQUEZ V L. Diagnosticity and multidimensional subjective workload ratings[J]. *Ergonomics*, 1996, 39(3):358-381.
- [19] BRUNKEN R, PLASS J L, LEUTNER D. Direct measurement of cognitive load in multimedia learning[J]. *Educational psychologist*, 2003, 38(1):53-61.
- [20] GAUVAIN M. Cognitive load theory[J]. *Psychology of learning & motivation*, 2010, 55(4):37-76.
- [21] BANHOLZER N, FEUERRIEGEL S, FLEISCH E, et al. Computer mouse movements as an indicator of work stress: longitudinal observational field study[J]. *Journal of medical internet research*, 2021, 23(4):1-12.
- [22] LURIA G, ROSENBLUM S. A computerized multidimensional measurement of mental workload via handwriting analysis[J]. *Behavior research methods*, 2012, 44(2):575-586.
- [23] JUST M A, CARPENTER P A. A theory of reading: from eye fixations to comprehension[J]. *Psychological review*, 1980, 87(4):329-354.
- [24] GOG T V, KESTER L, NIEVELSTEIN F, et al. Uncovering cognitive processes: different techniques that can contribute to cognitive load research and instruction[J]. *Computers in human behavior*, 2009, 25(2):325-331.
- [25] LIU H C, LAI M L, CHUANG H H. Using eye-tracking technology to investigate the redundant effect of multimedia web pages on viewers' cognitive processes[J]. *Computers in human behavior*, 2011, 27(6):2410-2417.
- [26] KRZYSZTOF K, 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.
- [27] KONTOGIORGOS D, J GUSTAFSON. Measuring collaboration load with pupillary responses—implications for the design of instructions in task-oriented HRI[J]. *Frontiers in psychology*, 2021, 12(7):1-19.
- [28] 郑玲,乔旭强,倪涛,杨威,李以农.基于多维信息特征分析的驾驶人认知负荷研究[J].*中国公路学报*,2021,34(4):240-250.
- [29] CASTRO-MENESES L J, KRUGER J L, DOHERTY S. Validating theta power as an objective measure of cognitive load in educational video[J]. *Educational technology research and development*, 2020, 68(2):181-202.
- [30] GHANI U, SIGNAL N, NIAZI I K, et al. ERP based measures of cognitive workload: a review [J]. *Neuroscience & biobehavioral reviews*, 2020, 118(7):18-26.
- [31] 李金波.网络学习环境中影响学习者认知负荷的因素[J].*电化教育研究*,2009(9):37-41.
- [32] WANG C, GUO J. A data-driven framework for learners' cognitive load detection using ECG-PPG physiological feature fusion and XGBoost classification[J]. *Procedia computer science*, 2019, 1479(1):338-348.
- [33] MERAT N, LAI F, JAMSON S L. The comparative merits of expert observation, subjective and objective data in determining the effects of in-vehicle information systems on driving performance[J]. *Safety science*, 2011, 49(2):172-177.

(下转第 54 页)