

多模态数据支持的教育科学研究发展脉络与挑战

彭红超 姜雨晴

【摘要】多模态数据凭借其全面、精准、保真地刻画学习画像的优势,成为支持教育科学研究的新兴趋势,但其二十余年的发展脉络并未得到系统探析。对此,本研究结合计量统计分析与内容分析方法,对209篇国内外核心期刊文献进行了全面、深入的解析,发现:国内外多模态数据支持的教育科学研究均大致经历了萌芽期、扩列期、裂变期三个阶段;在内容方面,国外研究呈多模态话语和多模态生理两个“小宇宙”态势,国内则集中于后者;在数据分析方面,国内外均呈多模态话语分析和多模态学习分析并驾齐驱之势;其理念发展经历了三方面的演变,在数据证据方面由行为证据扩展至生理证据,在数据分析方面由统计分析向多模态学习分析转变,在学习机制方面由以事件为中心向以人为中心转变;多模态数据融合、研究范式转变和数据隐私依然是当前面临的发展挑战。这些发现有利于研究者全面认识多模态数据在教学实践与科学研究中的作用、价值,精准把握其研究与发展的取向。

【关键词】多模态数据;学习分析;教育科学研究;数据融合;研究范式;研究走势;研究态势;研究挑战

【作者简介】彭红超,博士,助理研究员,华东师范大学开放教育学院(200062);姜雨晴,博士研究生,华南师范大学教育信息技术学院(510631)。

【原文出处】《中国远程教育》(京),2022.9.19~26,33

【基金项目】本文系中央高校基本科研业务费项目“智慧教育中教师与AI协同的教学决策策略研究”(2021ECNUYYJ030)的研究成果。

一、引言

教育数据已成为教育领域的巨大财富,正重塑着教学日常实践与科学研究设计。目前常用的教育数据除了作业、测验,还有在线学习平台中的学习日志、互动文本等,这些数据虽然已经能够描绘较为细致的学习者画像,但这仍然只是“侧面肖像”,难以准确揭示学习者学习规律。可喜的是,随着低成本、便携式的可穿戴设备在教育中的应用,教育数据已经开始扩展至脑电、心率、皮电等生理数据。这种多模态化的数据全方位、立体化地呈现了学习者的学习过程,为教育科学研究提供了更为丰富的数据源。研究发现,使用多模态数据分析学习者学习效果的准确率高出使用单一数据(Giannakos, Sharma, Pappas, Kostakos, & Velloso, 2019),并且能够为深入探究学习者行为的内在学习机理提供机会和支持(牟智佳, 2020)。多模态数据支持的教育科学研究是新兴发展趋势,但其事实上已有二十多年发展历史。厘清其研究走势与态势、理念的发展变化以及发展瓶颈等问题,有利于学者与教学实践者全面认识多模态数据,精准把握多模态数据

在教学实践与科学研究中的作用、价值与发展取向。

在目前已有的研究综述中,国内学者翟雪松等(2020)发现多模态数据中的生理数据主要用于支持监测情绪与控制、探究认知与控制、监测与改善健康三个方面,并探明了心率、呼吸、皮温、皮电等模态数据指标表征的意义;穆肃等(2021)发现用于学习分析的多模态数据有生理特征、心理测量、环境场景等五种,分析指标有行为、注意、投入等十余种,数据整合方式有“多对一”“多对多”“三角互证”三种。他们对多模态数据应用于教育研究的价值进行了挖掘,并对融合分析与解释等发展瓶颈进行了探索,但并未系统回答上述系列问题。对此,本研究综合运用计量统计分析与内容分析方法,通过深入解读中外核心期刊文献,探索这些问题的答案。

二、文献来源与分析

依据布拉德福文献离散规律,“大多数关键文献一般集中在少数核心期刊”(张斌贤,等,2009),本研究采用中文篇名“多模态数据、多模态+大数

据、多模态+学习分析、多模态交互分析、多模态话语分析”和英文篇名词“multimodal data、multimodal learning analytics、multimodal big data”分别在中国知网和 Web of Science 数据库核心合集检索 CSSCI 和 SSCI 两类期刊,检索时间为 2022 年 1 月 20 日。检索到的文献按以下标准遴选文献:①文献不能重复;②文献的主题需与教育科学研究有关。最终得到文献 209 篇,其中中文文献 81 篇,英文文献 128 篇。

首先,本研究采用文献计量法,对国内外多模态数据支持的教育科学研究走势进行分析。之后,将文献中的关键词导入 Bicom 2.02 软件,进行数据清洗与词频分析,并对词频大于等于 3(李运福,等,2018)的高频关键词进行词频共现分析。最后,将共现矩阵导入 Ucinet 6 软件,并使用其中的 NetDraw 工具进行社会网络中心度分析,得到可视化的研究态势图谱。同时,运用内容分析法深入解读研究内容,探析其研究态势、发展演变和面临的挑战。

三、研究热度与研究态势

(一)研究热度

计量统计分析结果如图 1 所示。由此可发现,多模态数据支持的教育科学研究最初并没有得到学者们的广泛关注,直到最近几年才逐渐引起学者们的探索兴趣。纵观其研究历程,国内外均有较大波动幅度,特别是国内,虽然起步晚,但发展迅速且具有明显的赶超之势。进一步内容分析还发现,国内外的研究热度大致相同,均经历了萌芽期、扩列期、裂变期三个阶段。

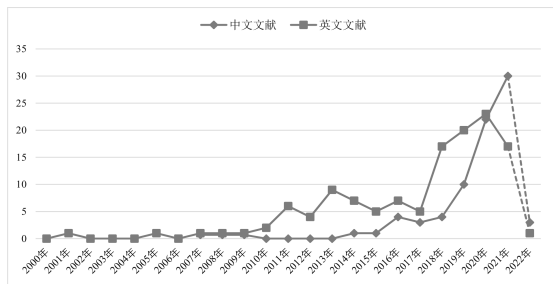


图 1 国内外多模态数据支持的教育科学研究文献分布

1. 萌芽期(国外 2000-2009 年,国内 2007-2013 年)

文献统计显示,伦敦大学学者凯里·朱伊特(Carey Jewitt)于 2000 年提出将多模态应用于语言教学,国外由此进入了多模态数据支持的教育科学研究萌芽期。语言学认为学习是多模态的,在学习过程中学习者需要依据学习兴趣和语境不断切换语言、视觉和动作等模态以获得知识与技能。数

据显示,直至 2007 年我国才开始有学者关注这方面的研究,标志性事件是学者顾曰国提出的较为公认的多模态学习行为研究假设,这一假设勾画了研究多模态学习的提纲,成为后续多模态话语分析的参照标准。

这一阶段的多模态主要是不同感官通道、语言、手势等社会意义表达的不同样态。有关这种多模态的研究,甚至深入到了诸如语言障碍、智力障碍、听力障碍、自闭症等特殊教育领域(Spinath, Wolf, Angleitner, Borkenau, & Riemann, 2005):通过自我报告、教师反馈、课堂观察等方式收集特殊儿童学习与交互过程中的多模态数据,运用多模态数据分析了解特殊儿童的学习状态与实时需求,并提供适应性支持以维持特殊儿童的认知和情感参与程度。统计数据显示,国内的萌芽期晚于国外,直到国外进入这一时期的尾端,国内才开始步入多模态研究时代。

2. 扩列期(国外 2010-2016 年,国内 2014-2016 年)

在萌芽期,学者们关注的感官通道只有学习者的听觉和视觉,并没有关注触觉等其他通道。直到 2010 年,阿姆斯特朗州立大学学者哈姆扎·卢普(Hamza-Lup)将触觉感官引入多模态领域,他认为学习者的学习应是五大感官相互作用的结果,其中触觉是极为重要的模态(Fernández-Cárdenas & Silveyra-De La Garza, 2010)。法国知觉现象学家梅洛-庞蒂(Merleau-Ponty)提出人类对世界的认知主要依靠具身经验,触觉是人类与外界物理环境交互所采用的最原初、最直接且最重要的感知系统(赵瑞斌,等,2021)。这种具身理论观点为触觉成为多模态数据重要组成部分提供了依据。由此多模态数据支持的教育科学研究也相继进入扩列期。

这一时期,国外学者在课堂中开展了一系列多模态数据交互行为研究,如师生、生生之间的语言、肢体等交互行为的研究,甚至有学者通过体育教学中的身体互动数据来探究学习者的课堂互动特征与模式。随着在线会议系统、网络平台、虚拟现实等技术在课堂中的应用,多模态数据的收集扩展至远程学习领域,如利用 Moodle、Skype、桌面视频会议和三维软件等工具开展的在线学习研究,开始注重收集手势、目光接触、语言交流、脸部头像和平台记录(如学习日志)等模态数据。数据显示,2014 年我国开始进入扩列期,标志性事件是郑瑶菲等学者(2014)在基于“云服务”的多模态课堂口头报告研究中将触觉、嗅觉、味觉等生命体感渠道纳入学生的多模态交互中。这一时期,我国持续时间很短,并和国外同一年进入裂

变期。

3. 裂变期(国外 2017 年至今,国内 2017 年至今)

随着低成本、便携式的可穿戴设备进入大众视野,支持教育教学研究的多模态数据不再局限于师生、生生的互动数据,以及平台记录的学习印记数据。表征学习者自身生理状态的数据也成为多模态数据的新兴成员。其标志性事件是,2017 年第七届学习分析与知识国际会议明确将融合生理数据的学习分析作为其全面拓展实践新取向的事实。其中缘由是,这种多模态分析能够提供更为精细的数字画像,做出更为精准的科学预测。有研究发现,单纯以学习行为数据(如按键、触屏、鼠标点击等)作为证据预测学习效果的错误率高达 39%,加入调查数据(如自我报告)后错误率可降至 15%,而结合生理数据(如脑电、心率、皮电等)后,错误率则可低至 6%(Giannakos, et al., 2019)。这种优势为多模态数据支持的教育科学研究带来了新机遇,也是促使国内外研究井喷式发展、进入裂变期的主要动力。

内容分析显示,国内外几乎同一时间采用生理数据开展研究(2017年)。这一时期,国内外学者开始关注使用脑电、心率、皮电等生理数据分析学习者在 学习过程中情感、认知和行为等方面的变化,研究侧重点也相应地向学习者内在学习机理转变。随着 VR 技术在教学中的广泛应用,学习环境也出现了由传统课堂向虚拟环境扩展的趋势,表现为国内外学者借助多模态数据追踪学习者在虚拟世界中的行为、情感和认知状态,探究 VR 技术对学习的促进作用。数据显示,受新冠疫情影响,国外研究在 2020 年出现了明显的拐点,而国内由于疫情的有效控制,其相关研究发展趋势并未发生明显变化,且在 2021 年出现了超越国外的局势。

(二) 研究态势

借助 Ucinet 6 软件中 NetDraw 工具进行社会网络中心度分析,形成国内外文献高频关键词社群图,如图 2 和图 3 所示。在此社群图中,方框越大则关键词的中心度越高,方框之间的距离越小则表示关键词之间越亲密(宋慧玲,等,2017)。

1. 研究关注的内容

观察图2和图3可发现,国内研究比国外研究更具有“凝聚力”,即各内容之间的关系更亲密。当然这并不表示国内研究比国外研究更精细。内容分析发现,国内研究大多关注理论的梳理与构建,但对其实际效用的验证相对缺乏。

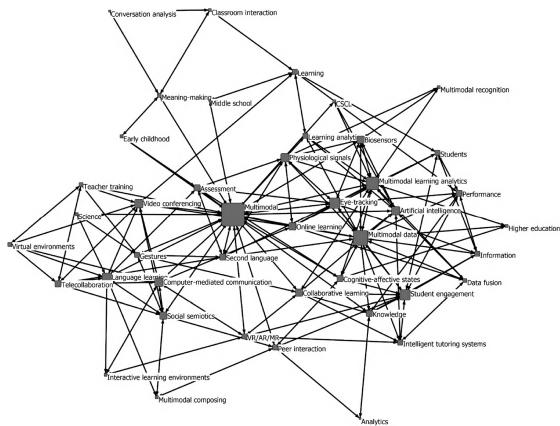


图2 国外文献高频关键词社群图

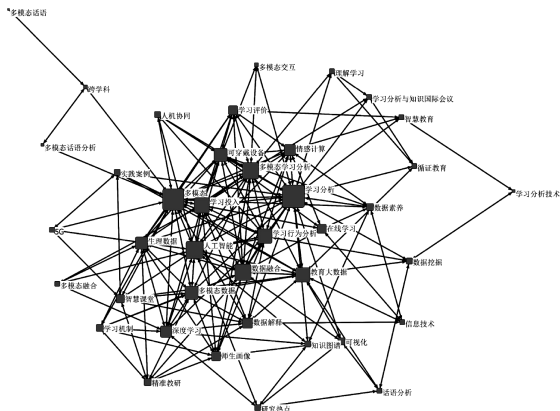


图3 国内文献高频关键词社群图

总体来看,国内外研究各具特色;国外研究呈两个“小宇宙”态势,一个是多模态话语,一个是多模态生理,而在国内后者是研究的核心问题,处于绝对中心位置。具体而言,国外在第一个“小宇宙”多模态话语方面,主要是以韩礼德(Halliday)的社会符号学为基础,基于师生在语言教学中互动媒介的变化对以师生互动为主的多模态话语分析开展大量研究,而在第二个“小宇宙”多模态生理方面,主要是利用多样的生理数据(特别是眼动数据),通过学习分析范式探究认知投入、情感投入等方面的学习内部机制。这方面的研究已发现,学习者的情感状态对于其认知、动机和调节过程均有影响,无论是积极还是消极的情感状态,均能够在学习过程中发挥重要作用(黄涛,等,2020; Ahn & Harley, 2020)。

国内研究和国外研究一样,发起于语言学领域,并且有较长的发展史,但直到现在仍仅是以顾曰国为代表的少数学术团队致力于此方面的研究,因此处在图5社群图的边缘地带。更有趣的是,国内其他领域学者之所以关注到此方面主要缘于多模态生理数据的影响,受此影响,国内学者在梳理

多模态起源时追溯到此。从图3社群图看,多模态生理数据支持教育科学研究处于中心地带,主要聚焦于多模态数据在学习行为、学习投入、学业情绪等方面的分析、解释与画像描绘,这方面国内学者尝试借助智能技术优势,关注智慧课堂环境,并追求深度学习。研究发现,通过人工智能及便携式可穿戴设备捕捉和测量学习者的课堂参与行为,可以准确预测学习者的学习状态(Sharma, Papamitsiou, & Giannakos, 2019),也能够为学者深入理解学习机制、发现学习规律提供可能(张琪,等,2019)。

另外,国外研究关注的对象较为广泛,涉及幼儿、儿童、中学生、大学生,甚至成人学习者(如教师培训)和特殊儿童。研究发现,通过多模态数据分析自闭症、智力障碍、语言障碍等学习者的突出学习行为变化(Spinath, et al., 2005; Jansson, 2011),可以为教师及时掌握其认知、行为、情感等参与状态提供重要参考(Moon & Sokolikj, 2020)。国内研究关注的对象主要是大学生群体,这与近年来国内加强人体实验伦理审查有关,特别是对中小学生的伦理保护。

2. 多模态数据分析方法

由图2和图3可知,在数据分析方面,国内外的态势大致相同,均涵盖两大类:多模态话语分析和多模态学习分析。这与国内外的多模态数据发展历程趋同有关。

内容分析发现,早期多模态话语分析的内容主要包括师生课堂对话,并主要局限于定性研究且多学科适应性较差。随着多媒体技术应用于课堂,其内容逐渐囊括了声音、动态图像、语料库等,同时定量分析成为主要研究范式,多学科视角的对话与融合成为关注焦点。尽管国内的多模态话语分析并未成为主流,但随着视频分析技术在教育研究中的应用,多模态话语分析也得到更多学者的青睐。

学习分析技术成为分析学习者学习流的主要范式得益于教育大数据的应用,学习分析的内容既有平台自动记录的线上学习数据(钟薇,等,2018),也有录像、声音、表情、手势等模态数据(Worsley, 2012)。可穿戴设备进一步丰富了可收集和分析的数据源,使得眼动、脑电、心率、皮电等生理数据在学习分析中得以关注(张琪,等,2016),学习者特征的刻画也更为全面、准确。目前,多模态学习分析主要有两个方面的作用:①为研究者提供证据;②为教师或学习者提供反馈。相比于前者,后者明显不同的是,生理数据多局限于心率、血压等智能腕表能够监测的数据,并且对于自动化分析的要求

更高。在这方面,现有智能课堂主要通过低水平甚至零水平的数据融合来提高反馈的及时性,尽管学者均一致认为综合分析多种模态数据是精准提供反馈与实时改进教学的前提(Bahreini, Naddski & Westera, 2016; Luo, Chen, Wang, & Liao, 2020)。内容分析发现,在为教师或学习者提供反馈方面,目前的研究主要集中在对学习者的情感的反馈,而对学习过程中认知、行为等方面的反馈未见有研究涉及。

四、研究的发展脉络与演变

通过文献梳理和内容分析发现,多模态数据支持的教育科学研究历经了三个方面的演变:①在数据证据方面,由行为证据扩展至生理证据;②在数据分析方面,由统计分析向多模态学习分析转变;③在学习机制方面,由以事件为中心向以人为中心转变。

(一)数据证据:由行为证据扩展至生理证据

语言学家认为学习是通过言语交流实现的,但也认为眼神、手势、动作等非言语交流是重要的知识“传播介质”,这些介质被认为是副语言因素。此时期的多模态数据证据主要是这些社会意义表达行为的模态。随着在线教学与管理平台的应用,特别是学习分析技术的加盟,在线学习的点击流数据模态成为新型行为证据。相应地,多模态数据支持的教育科学研究也由线下面对面课堂扩展至在线学习空间。以行为数据为证据开展教育科学研究,遵循一个基本假设,即心理构念是导致外在行为的原因,且外在行为可以推断心理构念(Kozak & Miller, 1982)。近年来,脑科学、神经科学等研究成果成为教育教学的重要理论支撑,特别是低成本、便携式可穿戴设备的应用,使得心理构念这一“黑盒”得以公之于众。相应地,多模态数据证据也由行为数据扩展至以脑电波为代表的生理数据。对于生理数据表征教育方面的心理构念,国内外学者已经进行了大量工作,如Jap等(2009)尝试利用 $(\alpha+\theta)/\beta$ 、 α/β 、 $(\alpha+\theta)/(\alpha+\beta)$ 、 θ/β 组合指标表征疲劳,翟雪松等人(2020)对心率、皮温、皮电等生理数据可以表征的心理状态做了系统梳理。生理数据“解剖”心理构念的优势已经得到教育学者的关注,它造就以生理数据为证据主体的多模态教育科学研究成为新的发展方向。

(二)数据分析:由统计分析向多模态学习分析转变

小数据时代,教育科学研究的多模态数据分析方法主要是统计分析,包括描述性统计、假设检验、相关分析、时间序列分析等。这种统计分析与单模态统计分析没有本质区别,只是不同模态数据如何

表征心理构念需要额外关注。在这方面,学者采用的方式主要有两种:第一,用不同模态数据同时从不同角度表征心理构念的整体情况,如巴塔查里亚等人(Bhattacharya, et al., 2021)使用面部表情(视觉方面)、语速语调(听觉)、文本语义三种模态数据分别表征情绪的整体水平;第二,用不同模态数据表征心理构念的各个维度情况,如沙玛等人(Sharma, et al., 2020)分别用眼动、面部表情、脑电等模态数据表征学生学习注意力、学业情绪、认知负荷等维度的强弱。

2015年,教育步入大数据时代,多模态数据分析也相应地向学习分析范式(数据清洗、特征提取、数据融合、结果报告)发展,这种分析涉及文本分析、话语分析、笔记分析、草图分析、情感分析、神经生理标记分析、眼动分析、多模态整合分析几大范畴(Blikstein, 2016)。支撑这些范畴分析的基础主要是算法模型,如社交网络模型、滞后序列分析模型、神经网络模型等。内容分析发现,多数学者主要是直接选用合适的算法模型开展相关教育研究,如胡钦太等人(2021)利用深度混合判别受限玻尔兹曼机神经网络模型来挖掘多模态学习行为的特征,并利用贝叶斯网络模型研究多模态数据与学习成绩的因果关系;曹晓明等人(2019)选用卷积神经网络中的ResNet 50模型来探索利用面部图像、脑电和学习日志等模态数据表征学习参与度。可喜的是,目前多模态学习分析的新发展取向已初见端倪,即学者们开始尝试探索算法模型的优化与创新。例如,马志强等人(2020)将社交网络与群体认知网络叠加,构建了一种社会认知网络模型来综合分析协作交互中的社会互动与认知互动水平;奥尔特加等人(Ortega, et al., 2019)在深度神经网络模型的基础上构建了一种新型模型,提高了利用视频、音频、文本等模态数据融合分析情感的精准度。

(三)学习机制:由以事件为中心向以人为中心转变

内容分析发现,在无线脑电仪器、眼镜式眼动仪等生物传感设备应用于教育科学研究以前,多模态数据分析主要以自我报告或观察到的行为数据为基础,这类数据的局限性导致教育科学研究只能探索以事件为中心的学习机制(Reimann, Markauskaite, & Bannert, 2014),即事件是不可分割的基本单元,学习机制通过事件之间的关系体现。从已有研究看,一般采用相关系数表征学习事件之间的相关或影响关系,采用序列图表征学习事件内隐的模

式与规律,采用社交网络表征事件主题的“亲密”关系。限制于技术水平,已有研究对学习事件的描述仅停留在描述性解释层面,且解释能力有限(Reimann, et al., 2014),也未对学习交互的内在学习机理进行深入探析。

以脑电、心电、皮电等生理数据在教育科学研究的应用为标志,多模态数据分析由以事件为中心发展至以人为中心。此时,学习事件这一“黑盒”得以打开,其内在在学习机理变得“可见”。相应地,研究内容也由关注学习行为、活动、结果等外在表现事件转向关注学习者的情感、认知、自我调控等内在在学习机理问题(尝试学习规律与特征由人的内在变化来表征)。内容分析发现,基于生理模态数据,学者们发现了诸多新的学习机理。例如,潘等人(Pan, et al., 2020)发现师生之间的“大脑共鸣”是教学策略有效的内在原因;梅苏拉姆(Meshulam, et al., 2021)、迪克尔(Dikker, et al., 2017)等诸多学者均发现,学生一班级的脑际耦合(学生个体与班级平均脑活动的相似性)是该学生学习表现的重要参考机制。虽然,这种以人为中心的多模态数据分析具有高度复杂性,但在生物传感器技术和大数据计算的支持下,其时间与人力成本已经表现出显著的降低趋势。

五、研究的挑战

由上述讨论可知,多模态数据支持的教育科学研究已经历了长足的发展与演变,不过,依然面临多模态数据融合、研究范式转变和数据隐私保护等挑战。

(一)多模态数据融合

互补性、时序性和层级性等特点造就多模态数据具有全面、精准、保真地表征学习者行为、认知、情绪等状态变化的潜能,而多模态数据融合是激发此潜能,实现探索学习者内隐心理变化和学习发生机制的关键,但这至少需要解决三个核心问题:①数据对齐或配准;②数据采样;③数据关联。对于第一个问题,内容分析发现,应用于教育科学研究的多模态数据采集设备多具有自动时间戳功能,这为各种模态数据在时间维度上的“对齐”提供了便利条件。但解决第二、第三个问题,仍面临许多挑战。第一,各种模态数据往往具有不可通约性(Non-commensurability),即不同模态数据的采集设备对不同现象的不同侧面敏感,导致各模态数据的物理单位不同且不可“通分”。在此情况下,如何实现不同模态数据的相互作用与启发是多模态数据融合

的首要任务 (Mechelen & Smilde, 2010)。第二,不同模态数据的分辨率往往不同,甚至非常悬殊,如脑电图 EEG 的时间分辨率为毫秒级,而功能磁共振成像 fMRI 的时间分辨率为秒级。在这样的情况下,如何在各模态数据中定位到有价值的信息是关键。第三,不同模态数据间既存在互补性信息,也存在不一致甚至冲突的信息。在这种情况下,如何去伪存真是发挥多模态数据优势必须要处理的问题。

不可通约性决定了在多模态数据融合时,首先需要确定适切的融合框架,这类框架主要有数据层融合、特征层融合、决策层融合三种。其中,数据层融合是对原始数据进行预处理后直接进行融合,当研究者收集的数据存在不可通约的情况时采用此融合框架通常比后两者框架更加复杂,因为研究者需要引入额外的信息“桥接”这些模态数据,而这类信息的确定是一项极具挑战性的工作。不同分辨率决定了在确定好融合框架后,还需完成数据采样,以便在压缩数据量的同时得到有价值的信息。纽厄尔 (Newell) 提出的认知带理论和安德森 (Anderson) 提出的带际间融合策略可作为采样依据,前者确定了生物带、认知带、理性带和社会带的时间尺度,这些尺度是定位“富有价值的信息”的关键区,后者确定了不同时间尺度的关联与拆解方案。不过,认知带理论以及带际间融合策略是否真的实用,还有待探究。对于多模态数据的互补与冲突,彭红超等 (2021) 提出的涵盖关联、叠加和拼接三种融合度的方案既顾及了去伪存真的解决路向,也提供了关联互补信息的思路。拉哈特等人 (Lahat, et al., 2015) 则提出了系列数学解决方案,包括解决去伪存真顺序处理模态思路以及互补信息的“软”“硬”链接方案。但这些构想的效用,同样需要经过长期的验证。

(二) 多模态研究范式转变

2007 年,计算机科学家吉姆·格雷 (Jim Gray) 提出“数据密集型科学发现”将成为继实验、理论、计算机仿真之后的第四科学研究范式 (Hey, Tansley, & Tolle, 2011), 强调通过仪器收集的各类大体量数据来认识与理解世界。多模态数据研究作为此范式在教育科学研究中的新代表,突破了学习过程“平面式描绘”的局限 (因数据主要为平台记录的行为印记使然), 转而走向“立体画像”的描绘 (由生理、心理、行为等数据使然) (祝智庭, 等, 2020); 突破了第一范式将人看作“关系之事本”, 第二范式将人看作“物之物本”, 第三范式将人看作

“实验之样本”的世界观 (米加宁, 等, 2018), 为落实第四范式的“对象之人本”提供了方法论。

不过,分析纳入研究的文献样本发现,以生理数据为主体的多模态数据研究并没有形成统一的研究规范供研究者们参考。第一,实验样本的体量依然参照针对小数据的统计分析体量标准,这是否满足多模态交互分析、多模态情感分析、多模态决策分析等新型分析方法的要求,还有待商榷。第二,多模态数据分析方法缺乏富有体系的前提条件检验准则,加之机器学习等智能算法分析的混入,导致研究结果可解性差的问题仍然没有有效的解决方法。第三,与一般研究不同,多模态研究的参与者需要“武装”更多种设备,会对被试造成更强烈的干扰,其霍桑效应也更明显,但如何消减此问题,未发现研究者涉及。第四,收集数据的设备与问卷等传统测量工具一样,也需要进行信效度检验,而这方面仍被研究者普遍忽视。第五,与一般量化研究采用提出假设—构建模型—设计实验—测量分析的自上而下的研究范式不同 (郑永和, 等, 2020), 多模态研究很多时候采用数据驱动的自下而上的研究范式,但这种范式的设计规范未见较有共识的范例。这些规范的缺失使得多模态研究的科学性、严谨性、有效性均无法精准评估,而这是教育科学研究向多模态研究范式转变不可回避的问题。

(三) 多模态数据隐私保护

多模态设备加之人工智能、5G 等新技术的应用,推动了教育数据全过程、立体式采集、分析的智能升级,但目前仍然存在较大的数据隐私泄露等风险。像学习者的点击流数据、学习印记等在线学习记录均存在网络泄露等安全隐患,另外皮电、心率和脑电等生理数据可在极大程度上揭示学习者的内在心理机制,也可在一定程度上反映学习者的身心健康状况 (Chassang, 2017), 虽然此类数据尚未达到生物医学研究水平,但依然是更为敏感的个人隐私,若被不法分子获取,会产生更为严重的不良后果。因此,多模态数据的隐私保护给数据安全带来更严峻的挑战,但这方面还未引起研究者足够的重视 (Beardsley, Martínez Moreno, Vujovic, Santos, & Hernández, 2020)。

对此,多模态数据支持的教育科学研究需要有数据安全与伦理的相关法律法规和从业规范的约束,并需要研究人员自身有较强的数据安全保护意识与能力。因为数据本身是中立的,研究者很难预估其收集的研究数据会对研究对象带来何种风险,

因此在开展研究前研究者有义务严格核对人体实验伦理审查的条目、完成相应的申请,并与实验对象及利益相关者签订知情同意书;研究收集的数据最好保存至本地磁盘后加密,而不转存至网盘中,并且设置存留期限,达成数据分析目的后及时清理;报告研究结果前,去除个人身份等敏感信息。更为重要的是,数据隐私保护需要多方协同努力制定相应的技术制度与标准,构建面向实践层面的数据脱敏机制(王一岩,等,2021),加强区块链等新技术支持的数据安全防护体系建设。这是消减数据隐私保护与数据价值应用之间矛盾(学生、家长一方面期望通过个人学习过程数据的分析获取全方位、精准化的学习支持服务,另一方面则会十分担忧个人数据隐私泄露(Ifenhaler & Schumacher, 2016))的关键措施。在这方面,我们还任重道远,并且已有的保护举措也存在不良反应,如知情同意书的引入明显降低了实验对象参与研究过程的程度(Beardsley, et al., 2020)。

六、结语

本研究通过对209篇相关研究文献分析发现,历经二十余年,多模态数据支持的教育科学研究经历了萌芽期、扩列期、裂变期三个阶段。目前,国外研究呈多模态话语和多模态生理两个“小宇宙”之势,国内主要集中于后者;在数据分析方面,国内外均呈多模态话语分析和多模态学习分析并驾齐驱之势。在这二十多年里,多模态数据支持的教育科学研究历经了三个方面的演变:①在数据证据方面,由行为证据扩展至生理证据;②在数据分析方面,由统计分析向多模态学习分析转变;③在学习机制方面,由以事件为中心向以人为中心转变。不过目前依然面临多模态数据融合、研究范式转变和数据隐私保护等方面的挑战。对此,本研究团队将持续致力于应对这些挑战的研究,特别是将聚焦研究多模态数据不可通约性的融合策略和数据驱动的自下而上的多模态研究范式设计。

参考文献:

- [1]曹晓明,张永和,潘萌,等.2019.人工智能视域下的学习参与度识别方法研究——基于一项多模态数据融合的深度学习实验分析[J].远程教育杂志,37(1):32-44.
- [2]胡钦木,伍文燕,冯广,等.2021.深度学习支持下多模态学习行为可解释性分析研究[J].电化教育研究,42(11):77-83.
- [3]黄涛,王一岩,张浩,等.2020.智能教育场域中的学习

者建模研究趋向[J].远程教育杂志,38(1):50-60.

[4]李运福,杨晓宏,周效章.2018.我国在线课程评价研究热点可视化分析与启示[J].中国远程教育,(7):70-78.

[5]马志强,管秀.2020.面向多维关联的社会认知网络分析——协作学习交互研究的新进展[J].远程教育杂志,38(6):96-103.

[6]米加宁,章昌平,李天宇,等.2018.第四研究范式:大数据驱动的社会科学研究转型[J].社会科学文摘,(4):20-22.

[7]牟智佳.2020.多模态学习分析:学习分析研究新增长点[J].电化教育研究,41(5):27-32,51.

[8]穆肃,崔萌,黄晓地.2021.全景透视多模态学习分析的数据整合方法[J].现代远程教育研究,33(1):26-37,48.

[9]彭红超,魏非,闫寒冰.2021.多模态数据赋能教师画像:从简笔画走向全息画像[J].开放教育研究,27(2):80-89.

[10]宋慧玲,张地珂,帅传敏,等.2017.我国网络学习绩效研究文献的可视化计量分析[J].中国远程教育,(3):30-36,74.

[11]王一岩,王杨春晓,郑永和.2021.多模态学习分析:“多模态”驱动的智能教育研究新趋向[J].中国电化教育,(3):88-96.

[12]张斌贤,陈瑶,祝贺,等.2009.近三十年我国教育知识来源的变迁——基于《教育研究》杂志论文引文的研究[J].教育研究,30(4):17-25.

[13]张琪,王红梅.2019.学习投入的多模态数据表征:支撑理论、研究框架与关键技术[J].电化教育研究,4(12):21-28.

[14]张琪,武法提.2016.学习分析中的生物数据表征——眼动与多模态技术应用前瞻[J].电化教育研究,(9):76-81.

[15]翟雪松,束永红,楚肖燕,等.2020.轻量级生理反馈技术的教育应用及测量——基于2015-2020年的文献综述[J].开放教育研究,26(6):37-46.

[16]赵瑞斌,张燕玲,范文翔,等.2021.智能技术支持下具身学习的特征、形态及应用[J].现代远程教育研究,33(6):55-63,83.

[17]郑瑶菲.2014.基于“云服务”的多模态课堂口头报告教学行动研究[J].中国电化教育,(12):133-138.

[18]郑永和,严晓梅,王晶莹,等.2020.计算教育论纲:立场、范式与体系[J].华东师范大学学报(教育科学版),(6):1-19.

[19]钟薇,李若晨,马晓玲,等.2018.学习分析技术发展趋向——多模态数据环境下的研究与探索[J].中国远程教育,(11):41-49,79-80.

[20]祝智庭,彭红超.2020.技术赋能智慧教育之实践路径[J].中国教育学报,(10):1-8.

[21]Ahn, B. T., & Harley, J. M. (2020). Facial expressions when learning with a Queer History App: Application of the control value theory of achievement emotions. British Journal of Educational Technology, 51(5), 1563-1576.

[22]Bahreini, K., Nadolski, R., & Westera, W. (2016). To-

wards multimodal emotion recognition in e-learning environments. *Interactive Learning Environments*, 24(3), 590–605.

[23] Beardsley, M., Martínez Moreno, J., Vujovic, M., Santos, P., & Hernández, L. D. (2020). Enhancing consent forms to support participant decision making in multimodal learning data research. *British Journal of Educational Technology*, 51(5), 1631–1652.

[24] Bhattacharya, P., Gupta, R. K., & Yang, Y. (2021). Exploring the contextual factors affecting multimodal emotion recognition in videos. *IEEE Transactions on Affective Computing*, 9, 135559–135570.

[25] Blikstein P. M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of Learning Analytics*, 3(2), 220–238.

[26] Chassang, G. (2017). The impact of the EU general data protection regulation on scientific research. *ecancermedical-science*, 11(709–727), 1–12.

[27] Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., et al. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, 27(9), 1375–1380.

[28] Fernández-Cárdenas, Juan Manuel, & Silveyra-De La Garza, Marcela Lucía. (2010). Disciplinary knowledge and gesturing in communicative events: A comparative study between lessons using interactive whiteboards and traditional whiteboards in Mexican schools. *Technology Pedagogy & Education*, 19(2), 173–193.

[29] Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management*, 48, 108–119.

[30] Hey, T., Tansley, S., & Tolle, K. (2011). The fourth paradigm: Data-intensive scientific discovery. *Proceedings of the IEEE*, 99(8), 1334–1337.

[31] Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, 64(5), 923–938.

[32] Jansson, A. B. (2011). Becoming a narrator: A case study in the dynamics of learning based on the theories/methods of Vygotsky. *Mind, Culture & Activity*, 18(1), 5–25.

[33] Jap, B. T., Lal, S., Fischer, P., & Bekiaris, E. (2009). Using EEG spectral components to assess algorithms for detecting fatigue. *Expert Systems with Applications An International Journal*, 36(2p1), 2352–2359.

[34] Kozak, M. J., & Miller, G. A. (1982). Hypothetical constructs versus intervening variables: A re-appraisal of the three-systems model of anxiety assessment. *Behavioral Assessment*, 4, 347–358.

[35] Lahat, D., Tülay Adal, & Jutten, C. (2015). Multimodal data fusion: An overview of methods, challenges and prospects. *Proceedings of the IEEE*, 103(9), 1449–1477.

[36] Luo, Z., Jingying, C., Guangshuai, W., & Mengyi, L. (2020). A three-dimensional model of student interest during learning using multimodal fusion with natural sensing technology. *Interactive Learning Environments*, (8), 1–14.

[37] Mechelen, I. V., & Smilde, A. K. (2010). A generic linked-mode decomposition model for data fusion. *Chemometrics & Intelligent Laboratory Systems*, 104(1), 83–94.

[38] Meshulam, M., Hasenfratz, L., Hillman, H., Liu, Y. F., Nguyen, M., Norman, K. A., et al. (2021). Neural alignment predicts learning outcomes in students taking an introduction to computer science course. *Nature Communications*, 12(1), 1922.

[39] Moon, J., Ke, F., & Sokolij, Z. (2020). Automatic assessment of cognitive and emotional states in virtual reality-based flexibility training for four adolescents with autism. *British Journal of Educational Technology*, 51(5), 1766–1784.

[40] Ortega, J., Senoussaoui, M., Granger, E., Pedersoli, M., Cardinal, P., & Koerich, A. L. (2019). Multimodal fusion with deep neural networks for audio-video emotion recognition Article 03196. Retrieved June 15, from <https://arxiv.org/abs/1907.03196>.

[41] Pan, Y., Dikker, S., Goldstein, P., Zhu, Y., & Hu, Y. (2020). Instructor-learner brain coupling discriminates between instructional approaches and predicts learning. *NeuroImage*, 211, Article 116657. Retrieved June 15, from <https://doi.org/10.1016/j.neuroimage.2020.116657>.

[42] Reimann, P., Markauskaite, L., & Bannert, M. (2014). e-Research and learning theory: What do sequence and process mining methods contribute?. *British Journal of Educational Technology*, 45(3), 528–540.

[43] Sharma, K., Papamitsiou, Z., Olsen, J. K., & Giannakos, M. (2020). Predicting learners' effortful behaviour in adaptive assessment using multimodal data [C]//10th International Learning Analytics and Knowledge (LAK).

[44] Sharma, K., Papamitsiou, Z., & Giannakos, M. (2019). Building pipelines for educational data using AI and multimodal analytics: A "grey-box" approach. *British Journal of Educational Technology*, 50(6), 3004–3031.

[45] Spinath, F. M., Wolf, H., Angleitner, A., Borkenau, P., & Riemann, R. (2005). Multimodal investigation of personality and cognitive ability. Results from two German twin studies bilsat and gosat. *Zeitschrift Für Soziologie Der Erziehung Und Socialisation*, 25(2), 146–161.

[46] Worsley, M. (2012). Multimodal learning analytics—Enabling the future of learning through multimodal data analysis and interfaces. *Acm International Conference on Multimodal Interaction*. ACM.