

# 基于人工智能引擎自动标注的课堂教学行为分析

卢国庆<sup>1</sup> [美]谢 魁<sup>2</sup> 刘清堂<sup>1</sup> 张臣文<sup>3</sup> 于 爽<sup>1</sup>

(1. 华中师范大学 湖北省教育信息化研究中心 湖北武汉 430079;

2. 俄亥俄州立大学 数字化学习研究实验室 美国俄亥俄州 43210;

3. 湖北文理学院 教育学院 湖北襄阳 441053)

**[摘要]** 课堂教学行为是影响课堂教学效果的重要因素,已有的课堂教学行为采集存在劳动密集、分类模糊和编码复杂等不足。人工智能技术为课堂教学行为大数据伴随式采集、自动化智能标注提供了新的契机。本研究以西北地区某市三所学校的1201个常规课堂教学视频为研究样本,利用人工智能引擎自动标注课堂教学行为,并采用相关性分析、主成分分析、非参数差异性检验等方法,对课堂教学行为类型、规律及差异性进行分析。研究发现:1) 课堂教学各类行为的出现频率不等且相差较大,其中,读写、讲授、巡视行为占比较大,生生互动、师生互动占比较小;2) 多数课堂教学行为之间具有关联性,其中,应答与生生互动之间的相关性最高,巡视、读写与其他行为存在负相关;3) 教师行为和学生行为并非完全属于不同的成分;4) 不同特征教师的课堂教学行为之间存在差异;5) 不同类型课堂的教学行为之间存在差异。研究结论可为人工智能时代挖掘课堂教学行为规律、改进课堂教学及开展教研活动提供参考。

**[关键词]** 人工智能; 课堂教学行为; 智能标注; 相关性分析; 差异性分析

**[中图分类号]** G424.1

**[文献标识码]** A

**[文章编号]** 1007-2179(2021)06-0097-11

## 一、引言

随着人工智能等技术的迅速发展,人类的生产生活方式正发生着重大变革。教育生态中教学要素的关系、课堂教学管理和评价方式也随着新技术的引入而发生重大变化。2017年7月,国务院(2017)

印发的《新一代人工智能发展规划》明确提出,利用智能技术加快推动教学方法、人才培养模式变革,开发智能教育助理,建立智能、快速、全面的教育分析系统,推动人工智能在教学、教研、管理等领域的应用。课堂是教育生态的主场,是教学行为最常发生的地方(王陆等,2019)。每种新技术都为课堂教学

**[收稿日期]** 2021-07-07

**[修回日期]** 2021-10-11

**[DOI 编码]** 10.13966/j.cnki.kfjyyj.2021.06.011

**[基金项目]** 国家自然科学基金“面向精准教学的课堂教学行为大数据分析及其应用策略研究”(71974073);“数据驱动的学习动机诊断模型及应用研究”(61977035);教育部人文社会科学研究规划基金“教师工作坊中学习者在线学习投入研究”(20YJA880078)。

**[作者简介]** 卢国庆,在读博士生,华中师范大学湖北省教育信息化研究中心,研究方向:数字化学习理论及实践研究(luguoqing@mails.ccnu.edu.cn);谢魁,美国俄亥俄州立大学教育学院西弗特杰出教授,华中师范大学长江学者讲座教授,研究方向:数字化学习、学习分析与建模研究(xie.359@osu.edu);刘清堂(通讯作者),教授,博士生导师,华中师范大学教育信息技术学院院长,研究方向:教师职业发展、学习分析与知识服务的研究(liuqtang@mail.ccnu.edu.cn);张臣文,副教授,湖北文理学院教育学院,研究方向:教师教育发展、信息化教学研究(309172202@qq.com);于爽,在读博士生,华中师范大学湖北省教育信息化研究中心,研究方向:学习分析(yushuang@mails.ccnu.edu.cn)。

**[引用信息]** 卢国庆,谢魁,刘清堂,张臣文,于爽(2021).基于人工智能引擎自动标注的课堂教学行为分析[J].开放教育研究,27(6):97-107.

带来了新的可能。人工智能技术能否创新、以何种方式创新课堂教学成为教育界关注的问题。

有效的课堂教学行为分析可以帮助教师了解课堂实情,改善课堂教学,促进教师专业发展(Mintzes, 2010)。目前,课堂教学行为实证分析方法主要有三种:第一,以“时间取样法”为主的量化分析(孙众等, 2020),即以课堂教学视频为研究对象,采用人工编码量化课堂教学言语和行为;第二,以“关键教学活动或事件”为主的质性分析(刘喆等, 2020),即以教学活动或关键事件为切入点(孙众等, 2013; 郁晓华等, 2018),运用内容分析法,借助质性分析软件对若干课堂教学视频样本进行编码,然后对编码结果进行统计、分析和归纳;第三,针对单一课堂视频语料的多模态话语分析(Lim, 2017; Lim, 2019),即通过分析课堂教学视频中言语、表情、手势、姿态等的符际关系,重点探索语言学、外语课堂教学中各模态之间的协同关系以及多模态的应用。然而,目前大多数课堂教学行为分析仍采用课堂观察(司治国等, 2012)、人工编码的方式(蒋立兵等, 2018)。这些传统方式受人工编码的限制,存在过度依赖专家、编码复杂、分析低效等问题(蒋立兵等, 2020),无法对大样本课堂教学视频进行自动标注分析。此外,由于数据标注技术的限制,研究者难以通过采集大样本数据分析课堂教学行为之间的关联关系和差异性。

人工智能在教育研究领域的逐步应用(余胜泉, 2018),形成了新兴的跨学科的研究领域——人工智能教育(Artificial Intelligence in Education, AIED)。尽管当前人工智能的整体发展仍处于弱人工智能阶段(刘凯等, 2018),但是人工智能为大规模采集课堂教学视频及自动智能分析提供了可能。人工智能和物联网技术构建的以智能设备为基础的数据采集体系,可实时捕获课堂中师生全方位的多模态学习行为数据,为刻画教师、学生课堂表现提供了新思路(顾小清等, 2021)。那么,目前课堂教学行为有哪些标注方式? 进展如何? 大规模课堂教学行为具有什么特征? 不同教师、课堂类型的教学行为有何差异? 本研究在深入梳理课堂教学行为采集方法演化和分析模型的基础上,采用人工智能引擎自动标注方法,以大量真实常规课堂的教学视频为研究样本,运用统计分析方法探索课堂教学行为之

间的关联特征,以及不同教师、不同课堂类型的教学行为差异。

## 二、文献综述

### (一) 课堂教学行为采集方式的进展

本研究将课堂教学行为的标注方式分为纯手工标注、半自动化标注、智能标注三个阶段(见图1),三个阶段按时间先后出现,并非替代性发展,而是同时存在的。

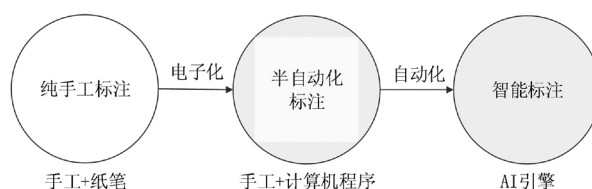


图1 课堂教学行为采集方式的发展历程

#### 1. 课堂教学行为的纯手工标注

传统的课堂教学行为采集主要采用纯手工标注,研究者通过现场观察,将课堂教学行为类别填入纸质记录卡片。早期的课堂观察与记录工具有座位表观察记录、弗兰德斯互动分析记录表、S-T分析表。座位表观察记录(seating chart observation records, SCORE)主要指课堂观察人员利用教室的座位表,依照预定的符号和编码框架,记录师生的课堂互动和移动状态(吴权威等, 2020)。早期的弗兰德斯互动分析系统每隔3秒对言语活动事件进行采样,全部取样后利用纸笔将事件的迁移序列记录在10×10的矩阵中,并统计课堂活动事件出现的频次(宁虹等, 2003)。S-T分析主要记录师生的课堂教学行为类别,每隔固定时间(如30秒)记录一次,根据该时刻课堂教学的行为类别,手工记录T行为或者S行为,从而判断行为转化和教学模式(程云等, 2015)。这些单纯依靠手工标注的方法存在分析维度少、劳动密集、容易出错等缺点。

#### 2. 课堂教学行为的半自动化标注

随着计算机技术的发展,辅助课堂教学行为标注和分析的软件不断涌现,可用于课堂教学行为的半自动化采集。常见的标注软件有LessonNote、Nvivo和MMA video 2.0等。LessonNote是一款保留了课堂座位表特征,通过画画、笔记、录音、照相等方式,标注师生互动行为的iPad应用程序(Lesson

Study Alliance 2020)。Nvivo 是一款打点标注、分析课堂教学视频中行为、教学事件等的质性分析软件。课堂多模态话语分析以教学录像为主要研究对象, 对课堂教学中语言、手势、媒体等符号特征进行标注。例如, 研究者通过多模态视频分析软件 MMA video 2.0 从多个层面对课堂录像进行标注、统计和分析(郭红伟等 2020; 陈松菁 2014)。ELAN 软件也是标注口语录音、教学录像等线性语料的常用软件( Lausberg & Sloetjes 2009)。这种半自动化标注的方法同样依靠人工打点标注, 很难实现大规模课堂教学视频中教学行为的即时采集与分析。

### 3. AI 支持的课堂教学行为智能标注

随着图像识别技术、机器学习算法、传感器技术的发展, 课堂教学行为的自动化、智能化标注与分析成为可能。研究者利用多种智能算法, 识别、提取及分析教学视频帧图像主体的行为特征。例如, 任婕(2019)通过自动识别课堂录像中的六类行为特征动作, 并利用多模态信息融合及层次分析法计算兴趣度。周鹏霄等(2018)通过检测视频帧图像轮廓特征、主体动作幅度等, 使用贝叶斯因果网模型获取行为序列, 智能识别课堂教学视频中的师生行为。刘清堂等(2019)基于课堂 S-T 行为分析理论, 从数据采集与存储、行为建模与计算、智能服务三方面提出课堂教学行为智能分析模型, 并通过实例进行分析和应用。赵刚等(2020)针对传统教学行为存在的手工编码、定时分隔等局限, 构建了多维度特征融合的教学行为智能分析模式。另外, 有研究者借助穿戴式便携传感器, 智能记录课堂教学行为。例如, 普列托等( Prieto et al., 2016)利用头戴式便携眼动仪、脑波仪、智能手机, 自动采集教师的眼动、脑波、三维加速度、视野录像和语音等数据, 使用机器学习算法预测教师发布任务、提问、解释等教学活动。然而, 已有算法的分析粒度、识别精度还有待完善。

AI 支持的课堂教学行为智能采集工具适用于不同的应用场景, 目前主要应用于智慧教室和常规教学。针对智慧教室情境, 研究者设计开发了苏格拉底教学行为分析系统。该系统根据师生课堂智能终端的使用情况, 智能标注师生教与学行为, 采集内容主要涉及技术运用频次、累计时间、互动指数、教法应用指数、技术运用分布图等( Ku et al., 2018;

吴权威等 2020)。此外, 智课系统是一款 AI 课堂教学行为分析系统, 主要利用人工智能、大数据、互联网、音视频等技术, 实现常规教学行为的常态化伴随式数据采集。该系统根据既定的课堂教学行为分类编码, 采用固定时间间隔(如 3 秒)的取样方式, 利用深度图像识别技术, 自动智能标注课堂教学行为的最佳类别编码, 生成课堂教学行为的结构矩阵, 然后通过对课堂教学视频的结构化分析, 借助数据挖掘技术对课堂教学行为进行建模, 生成可视化分析报告, 实现及时的评测和诊断。

三个阶段的课堂教学行为采集特征不同(见表一)。与纯手工标注、半自动化标注相比, 智能标注的识别精度尽管存在误差, 但为大规模处理课堂教学视频提供了技术支持。据此, 本研究主要采用人工智能引擎自动标注常规课堂教学行为, 对课堂教学行为的规律和差异进行量化分析。

### (二) 课堂教学行为分析模型

课堂教学行为分类、建模等相关研究由来已久, 且近年来备受瞩目。诞生于 20 世纪 60 年代的弗兰德斯互动分析系统( Flanders, 1963) 是一类影响较大的课堂行为分类系统。该分类模型将课堂言行为分为 10 类, 其中, 教师语言行为占 7 类, 学生语言行为占 2 类, 最后一类是沉默和混乱。同一时期, VICS 语言互动分类系统( 杨平展等 2009) 将课堂教学行为分为五类。为了减少行为分类的模糊性, 有研究者提出了 S-T 分析, 将 11 类课堂教学行为分为教师( T) 行为和学生( S) 行为两类( 傅德荣等, 2011)。

后续研究者针对不同的环境和学科提出了各种分类模型。例如, 顾小清等(2004)提出了基于信息技术的互动分析框架; 方海光等(2012)针对数字化课堂教学环境提出了改进型弗兰德斯互动分析系统( iFIAS); 张屹等(2016)提出智慧教室环境下的课堂教学互动分析编码系统; 刘向永等(2018)针对英语学科的特点, 修改和完善了 iFIAS, 提出了基于电子书包的小学英语课堂师生互动分析编码系统; 何文涛等(2018; 2019)对协作学习情境中的教学系统进行建模, 运用教学信息集合 IIS 结构, 将教学视频转录成文本, 并对信息流进行切分, 提出了基于视频的教学机制图。研究者发现, 课堂教学中 82% 的信息通过教师的非言语行为传递, 只有 18% 的信息通

表一 课堂教学行为采集方式的三个阶段

采集方式	常见工具	采集内容	优点	缺点
纯手工标注 (手工+纸笔)	座位观察记录表	师生互动、师生移动状态	直观形象	记录维度少
	课堂互动分析矩阵	传统教室中师生言语内容分析	将质性、量化研究融合, 影响较大, 便于教学反思	劳动密集、编码复杂、忽视非言语行为
	S-T 分析表	教师行为和学生行为	减少了分类的模糊性	师生行为可能同时成对出现
半自动化标注 (手工+计算机)	LessonNote	通过画画、笔记、录音、照相等方式记录师生互动	保留了座位表的特性, 记录内容丰富	劳动密集
	Nvivo	预定的编码框架	质性分析软件	劳动密集
	MMA video	用于标记和统计语言特征、手势、媒体等多模态信息	分析课堂教学中的多模态信息	劳动密集
	教学过程机制图	将教学视频转录成文本、确定角色行为编码体系、编码行为及归类角色、切分信息流	重点关注视频中的言语文本信息	忽视非言语行为
智能标注	苏格拉底教学分析系统	技术运用频次、累计时间、技术互动指数、技术运用分布图、教法应用指数	智慧教室技术运用情况	需配合智能终端使用, 应用场景受限
	智课系统	采集师生课堂教学行为: 学生行为包括举手、应答、听讲、读写、生生互动, 教师行为包括板书、巡视、讲授、师生互动 <sup>①</sup>	自动化大规模采集	识别精度存在一定误差

过言语行为传递(Brown, 2020)。由此可见, 课堂教学行为分析不仅需要考虑课堂的言语行为, 更需要关注课堂的非言语行为。

针对已有行为类别划分的模糊性与不均衡, 程云等(2017)从行为活动和行为主体两个维度, 将课堂教学行为分为学生言语行为、学生活动行为、教师言语行为、教师活动行为四类, 并建立了包含16类教学行为的云坐标体系, 提出了课堂教学行为云模型。本研究基于课堂教学行为云模型, 关注课堂的非言语行为, 结合图像识别技术、行为特征检测技术, 重点探索课堂教学过程中的教学行为特征和差异性规律。

### (三) 课堂教学行为特征及差异

教学行为指师生一系列有目的的教与学活动, 具有一定的主体性、关联性和可变性。从参与主体看, 课堂教学行为包括教师的教学行为和学生的学习行为; 从师生交互看, 课堂教学行为分教师主导为主、学生主导为主两类。从只关注教师行为和学生行为的S-T分析(傅德荣等, 2011), 到将课堂对话内容分为十类的弗兰德斯互动分析系统(Flanders, 1963), 再到涵盖16类课堂教学行为的云模型(程云等, 2017), 不同的理论分析框架将课堂教学行为划分为不同的类别, 为课堂教学行为的量化分析提供了模型。运用这些模型和描述性统计方法, 研究

者能够有效地分析课堂教学行为的频次或占比、时序变化、师生行为的主导趋势等。不同的课堂教学行为之间具有一定的关联性, 即某种教学行为会伴随着其他行为而出现。

研究表明, 不同特征教师群体的课堂教学行为有一定差异, 不同课堂类型的教学行为同样会表现出一定的特征差异。傅德荣等(2011)根据教师行为占比率、师生行为转换率, 将课堂教学分为练习型、混合型、讲授型和对话型等四类, 其中练习型课堂的教师行为占比低于30%, 讲授型课堂的教师行为占比高于70%, 对话型课堂的师生行为转换率大于40%, 混合型课堂同时满足教师行为占比大于30%、小于70%, 且师生行为转换率小于40%。王陆等(2020)发现, 随着时间的变化, 混合型、对话型课堂中的有效提问具有相似的变化趋势。然而, 由于数据采集技术的限制, 已有研究没有通过大规模量化研究分析课堂教学行为之间的关联特征, 且鲜有研究关注不同课堂教学行为的关联特征、不同教师特征及不同课堂类型的教学行为差异。

## 三、数据采集与分析方法

本研究基于课堂教学行为云模型, 以常规课堂教学视频为研究样本, 采用AI课堂教学行为分析系统对每堂课的教学行为进行自动化采集和标注。

### (一) 研究对象与样本

本研究于2019年9-11月采集了3所中学(1所高中, 2所初中)的1201份课例数据, 经过人工初步筛选, 获得976份有效课堂数据样本, 有效率为81.27%。在有效样本中, 高中课例有590份、初中课例有386份, 涉及语文、数学、英语等10个学科。其中, 数学(21.11%)、语文(18.65%)和英语(15.98%)的有效课例最多, 美术(0.1%)、心理(0.41%)、政治(1.84%)等学科的有效课例较少。女教师的课例(622份, 占63.73%)比男教师(354份, 占36.27%)多。高教龄组教师课例(579份, 占59.32%)比低教龄组教师(397份, 占40.68%)多。高级职称教师课例有397份, 占40.68%, 一级职称教师课例446份, 占45.70%, 二级职称教师课例133份, 占13.62%。课堂类型涉及4种, 包括讲授型课堂(137份, 占14.04%)、练习型课堂(457份, 占46.82%)、混合型课堂(25份, 占2.56%)、对话型课堂(357份, 占36.58%)。

### (二) 数据收集方法

本研究采用人工智能引擎自动标注、课堂视频观察和线下数据收集等三种数据收集方法。人工智能引擎自动标注主要依托课堂中搭建的人工智能终端, 对教学过程进行无干扰、伴随式常态化数据采集。在班级建模后, 研究者通过多视角(教师/学生视角)4K摄像机, 实时采集课堂教学视频信息, 经由人工智能引擎终端的边缘计算模块(见图2), 实

现对客户端视频数据的结构化数据处理, 并可视化呈现行为分布、教学模式等数据分析报告。

人工智能引擎首先实时提取课堂样本中的图像、语音/文本特征, 获得行为特征矩阵, 并经过深度神经网络的教师行为预测模型, 计算得到九类课堂行为的时间序列矩阵。该人工智能引擎能够自动识别课堂教学行为云模型中的九类行为, 其中, 学生行为包括举手、应答、听讲、读写和生生互动, 教师行为包括板书、巡视、讲授和师生互动。对于学生行为, 一个学生举手就被识别为举手行为, 应答指1个及以上学生站立回答问题, 听讲指大部分学生端坐、抬头和听课, 读写指多数学生在课堂低头阅读和书写, 生生互动指多个学生应答或者小组讨论。对于教师行为, 板书指教师在黑板、白板或大屏上书写, 巡视指教师在学生区域走动、观察或授课, 讲授指教师在讲台区域授课、讲解或演示, 师生互动涉及教师提问、学生应答、学生在讲台区域互动。教师授课时, 系统自动录制教学视频, 采用时间间隔采样的方法, 根据既定的课堂教学行为分类框架, 记录、描述和自动统计教师行为和学生行为。经过多轮大样本的图像训练和人工检验, 课堂教学行为为识别的精度大于80%。

课堂视频观察的主要任务, 是对人工智能引擎自动标注的结果进行人工观察和筛选, 排除无效课例。线下数据收集主要指从教育管理部门获取视频对应教师的个人信息, 如性别、职称、教龄等。

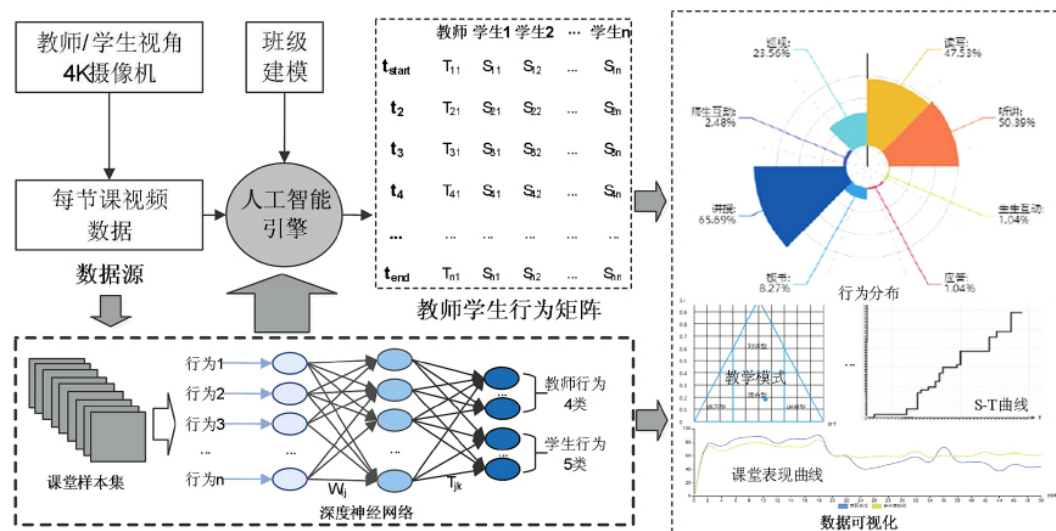


图2 人工智能引擎机制

### (三) 研究变量

研究变量包括课堂教学行为类型、教师个人信息、课堂类型三个维度,共13个研究变量。课堂教学行为类型变量用于衡量课堂教学行为的频率,由AI引擎自动标注获得,共包括9个变量,其中,学生行为包括举手、应答、听讲、读写、生生互动五个变量,教师行为包括板书、巡视、讲授、师生互动四个变量。教师个人信息包括教师性别、教龄、职称三个变量。根据教师行为占比、师生行为转换率,课堂类型变量被分为练习型、混合型、讲授型和对话型四种教学模式。

### (四) 数据分析方法

由于课堂行为数据呈非正态分布,为了更加科学准确地描述课堂各个行为之间的关系,本研究先采用斯皮尔曼相关系数对行为间关系进行相关分析,挖掘课堂教学行为之间的关联;其次,采用主成分分析方法将高维数据进行降维,以发现课堂教学行为的主成分特征;再次,采用非参数差异性分析方法分析不同教师特征的课堂教学行为差异;最后,分析不同教学模式的课堂教学行为之间的差异,并以课堂讲授为基准,采用非参检验方法分析不同课堂类型的教学行为差异。

## 四、数据分析与结果

### (一) 课堂教学行为关联分析

本研究采用斯皮尔曼相关系数度量非正态数据变量间的相关性(见表二)。由表二第二栏均值和标准差的描述性统计分析结果可知,课堂教学行为按照频率从高到低排列为:读写( $M = 1.08$ ,  $SD = 5.81$ )、讲授( $M = 0.92$ ,  $SD = 4.76$ )、巡视( $M = 0.58$ ,  $SD = 3.13$ )、听讲( $M = 0.26$ ,  $SD = 0.78$ )、应答( $M =$

$0.18$ ,  $SD = 1.41$ )、举手( $M = 0.06$ ,  $SD = 0.27$ )、板书( $M = 0.09$ ,  $SD = 0.30$ )、生生互动( $M = 0.03$ ,  $SD = 0.25$ )、师生互动( $M = 0.03$ ,  $SD = 0.12$ )。总体来看,常态化课堂教学过程中,师生互动和生生互动相对较少。根据标准差排名,课堂教学行为波动较大的三类行为分别是读写、讲授和巡视,波动较小的三类行为分别是举手、生生互动和师生互动。

表二中的相关性统计结果表明,多数课堂教学行为之间具有关联性。其中,教师巡视、学生读写与其他课堂行为均存在显著的负相关。此外,学生的应答行为与生生互动的相关性最高( $r = 0.59$ ,  $p < 0.01$ )。由此,生生互动可能会在一定程度上促进学生的应答行为,教师可以通过课堂教学活动设计促进生生互动、师生互动,进而提高学生应答行为发生的频率。

### (二) 课堂行为的主成分分析

针对976条课堂教学行为数据,本研究运用主成分分析法,按照因子载荷从大到小排序,并删除因子载荷小于0.6的教学行为,从九种课堂行为数据中提取出两个主成分,这两个主成分累积解释方差达到总方差的79.07%,说明其在79.07%的程度上代表九个维度的课堂行为。换句话说,所提取的两个主成分可以代表九种课堂教学行为,有效地实现了对课堂行为数据的降维处理。

结果表明:1)与主成分1相关性较高的行为依次是读写、巡视、讲授、生生互动、听讲和师生互动,载荷分别为0.918、0.849、0.815、0.795、0.775、0.642;2)与主成分2相关性较高的行为是举手、应答和板书,载荷分别为0.933、0.878、0.826。由此可见,在一定程度上,教学是一个有机的系统,课堂

表二 斯皮尔曼相关分析结果( $n = 976$ )

	均值(标准差)	举手	应答	听讲	读写	生生互动	板书	巡视	讲授
举手	0.06(0.27)	1							
应答	0.18(1.41)	0.04	1						
听讲	0.26(0.78)	0.05	0.16**	1					
读写	1.08(5.81)	-0.03**	-0.44**	-0.73**	1				
生生互动	0.03(0.25)	0.05	0.59**	0.10**	-0.30**	1			
板书	0.09(0.30)	0.11**	0.13**	0.11**	-0.04	-0.02	1		
巡视	0.58(3.13)	-0.09**	-0.22**	-0.28**	0.37**	-0.04	-0.24**	1	
讲授	0.92(4.76)	0.02	0.26**	0.29**	-0.30**	0.11**	0.01	-0.84**	1
师生互动	0.03(0.12)	0.14**	0.15**	0.14**	-0.20**	0.07*	-0.01	-0.21**	0.13**

注: \* 在0.05级别(双尾)相关性显著, \*\* 在0.01级别(双尾)相关性显著。

教学的教师行为和学生行为密不可分、相互联系,且有时候能够组成教学行为对,如读写和巡视、生生互动和师生互动。同时,教师行为和学生行为又具有一定的独立性,例如,学生举手、应答等是相对独立的行为。

### (三) 不同特征教师的课堂教学行为差异

为了进一步对课堂数据进行分析,本研究对教师的性别、教龄、职称等特征进行差异性分析。鉴于因变量均为非正态定距数据,本研究采用曼-惠特尼 U 检验(Mann-Whitney U-test)做差异分析。曼-惠特尼 U 检验又称曼-惠特尼秩和检验,可以看作是对两均值之差的参数检验方式的 T 检验或大样本正态检验的替代方法。

#### 1. 不同性别教师的课堂行为差异

在总样本中,女教师有 622 条记录,占 63.73%,男教师有 354 条记录,占 36.27%。通过 Mann-Whitney U 检验,本研究比较了男女教师的课堂行为。结果表明,与男教师相比,女教师课堂具有更多的举手( $Z = -5.281, P = 0.000 < 0.001$ )、应答( $Z = -8.095, P = 0.000 < 0.05$ )、听讲( $Z = -3.596, P = 0.000 < 0.05$ )、生生互动( $Z = -5.456, P = 0.000 < 0.001$ )、讲授( $Z = -7.758, P = 0.000 < 0.001$ )、师生互动( $Z = -2.676, P = 0.000 < 0.001$ )等教学行为;男教师课堂有更多的读写( $Z = -8.935, P = 0.000 < 0.001$ )和巡视( $Z = -8.157, P = 0.000 < 0.001$ )行为。男女教师的板书行为( $P = 0.11 > 0.05$ )不存在统计学意义上的差异。

#### 2. 不同教龄教师的课堂行为差异

在总样本中,低教龄组的教师有 397 条记录,占 40.68%,高教龄组教师有 579 条记录,占 59.32%。

通过 Mann-Whitney U 检验,本研究比较了高低教龄组教师的课堂教学行为差异(见表三)。统计结果表明,低教龄组教师的课堂有更多的巡视行为( $Z = -4.504, P = 0.000 < 0.001$ )和学生听讲行为( $Z = -2.857, P = 0.004 < 0.01$ ),高教龄组教师的课堂有更多的讲授行为( $Z = -5.299, P = 0.000 < 0.001$ )。除了听讲、巡视和讲授外,其余课堂教学行为在教龄上不存在显著差异。

#### 3. 不同职称教师的课堂行为差异

由于样本不服从正态分布,本研究采用鲁斯卡尔-沃利斯检验(Kruskal-Wallis test)检验高级、一级、二级职称教师的课堂行为是否存在差异。在总样本中,高级职称教师有 397 条记录,占 40.68%,一级职称教师有 446 条记录,占 45.70%,二级职称教师有 133 条记录,占 13.62%。统计结果表明(见表四),举手( $\chi^2 = 15.60, df = 2, p < 0.001$ )、读写( $\chi^2 = 9.70, df = 2, p < 0.001$ )、生生互动( $\chi^2 = 30.27, df = 2, p < 0.001$ )、板书( $\chi^2 = 24.70, df = 2, p < 0.001$ )、巡视( $\chi^2 = 42.07, df = 2, p < 0.001$ )、讲授( $\chi^2 = 30.08, df = 2, p < 0.001$ )等行为存在显著差异。此外,高级职称组教师的课堂有更多的举手、讲授行为,一级职称组教师的课堂有更多板书、读写行为,二级职称组教师的课堂有更多的巡视、生生互动行为。

#### (四) 不同课型的课堂教学行为差异

本研究先采用鲁斯卡尔-沃利斯检验,检验讲授型、练习型、混合型和对话型课堂是否在课堂行为上存在差异。统计结果表明,举手( $\chi^2 = 28.059, df = 3, p < 0.001$ )、应答( $\chi^2 = 35.498, df = 3, p < 0.001$ )、

表三 不同教龄教师的课堂行为差异统计

	低教龄组(n=397)		高教龄组(n=579)		曼-惠特尼 U 值	Z 值	显著性
	秩均值	秩和	秩均值	秩和			
举手	368.35	67776	367.22	201969	50444	-0.06	0.950
应答	347.1	63867	374.32	205878	46847	-1.50	0.132
听讲	406.16	74734	354.57	195011	43486	-2.86	0.004
读写	356.73	65638.50	371.10	204106.5	48618.50	-0.80	0.426
生生互动	358.70	66000.50	369.78	203010.5	48980.50	-0.62	0.535
板书	344.99	63478	375.03	206267	46458	-1.66	0.096
巡视	428.44	78832.50	347.11	347.11	39387.50	-4.50	0.000
讲授	295.80	54428	391.49	215317	37408	-5.30	0.000
师生互动	360.63	66356.50	369.80	203388.5	49336.50	-0.51	0.608



表四 不同职称教师的课堂行为差异统计

课堂行为	三类职称教师的均值			鲁斯卡尔-沃利斯检验	$\chi^2(df=2)$	有无显著性差异
	高级组	一级组	二级组			
举手	494.26	434.38	407.32	0.000***	15.596	有
应答	462.30	449.79	458.46	0.800	0.448	无
听讲	467.46	419.71	529.45	0.000***	18.342	有
读写	439.95	487.28	417.49	0.008***	9.699	有
生生互动	482.00	403.09	529.94	0.000***	30.266	有
板书	464.79	483.65	353.75	0.000***	24.702	有
巡视	404.42	470.48	571.79	0.000***	42.071	有
讲授	504.91	436.42	369.66	0.000***	30.081	有
师生互动	469.68	440.58	462.9	0.28	2.516	无

听讲( $\chi^2 = 653.907, df = 3, p < 0.001$ )、读写( $\chi^2 = 479.996, df = 3, p < 0.001$ )、生生互动( $\chi^2 = 29.985, df = 3, p < 0.001$ )、板书( $\chi^2 = 25.393, df = 3, p < 0.001$ )、巡视( $\chi^2 = 112.954, df = 3, p < 0.001$ )、讲授( $\chi^2 = 62.965, df = 3, p < 0.001$ )、师生互动( $\chi^2 = 31.853, df = 3, p < 0.001$ )等课堂教学行为,在不同课堂类型之间均存在显著差异。为进一步分析不同课堂类型之间的差异,本研究采用 Mann-Whitney U 检验法,将讲授型课堂分别与其他三类课堂进行差异检验。

#### 1. 讲授型课堂与练习型课堂的教学行为差异

与讲授型课堂相比,练习型课堂存在更高频率的读写( $Z = -16.66, P = 0.000 < 0.001$ )、巡视行为( $Z = -9.25, P = 0.000 < 0.001$ ),且存在显著差异。举手( $Z = -4.07, P = 0.000 < 0.001$ )、应答( $Z = -4.33, P = 0.000 < 0.001$ )、听讲( $Z = -17.18, P = 0.000 < 0.05$ )、生生互动( $Z = -3.45, P = 0.000 < 0.001$ )、讲授( $Z = -7.18, P = 0.000 < 0.001$ )、师生互动( $Z = -4.33, P = 0.000 < 0.001$ )等教学行为在讲授型课堂出现的频率更高,且存在显著差异。另外,在板书行为( $P = 0.38 > 0.05$ )方面,讲授型课堂和练习型课堂之间不存在显著差异。由此可见,与练习型课堂相比,讲授型课堂具有更高频率的举手、应答、听讲、互动等行为。

#### 2. 讲授型课堂与混合型课堂的教学行为差异

与讲授型课堂相比,混合型课堂存在较多的读写( $Z = -5.95, P = 0.000 < 0.001$ )、较少的听讲( $Z = -7.38, P = 0.000 < 0.001$ )和讲授( $Z = -2.46, P = 0.01 < 0.05$ )行为。但两种课堂类型在举手( $P =$

$0.35 > 0.05$ )、应答( $P = 0.66 > 0.05$ )、生生互动( $P = 0.97 > 0.05$ )、板书( $P = 0.38 > 0.05$ )、巡视( $P = 0.07 > 0.05$ )、师生互动( $P = 0.32 > 0.05$ )等行为方面不存在显著差异。两类课堂在很多行为方面没有表现出统计学意义上的差异,可能是因为混合型课堂( $n = 25$ )的样本较少。后续研究可以扩大样本量,进一步研究两类课堂教学行为之间的差异。

#### 3. 讲授型课堂与对话型课堂的教学行为差异

与讲授型课堂相比,对话型课堂存在更多的读写( $Z = -13.52, P = 0.000 < 0.05$ )、板书( $Z = -2.50, P = 0.01 < 0.05$ )和巡视( $Z = -4.10, P = 0.000 < 0.05$ )行为。然而,与讲授型课堂相比,对话型课堂的举手( $Z = -2.20, P = 0.03 < 0.05$ )、听讲( $Z = -13.67, P = 0.000 < 0.05$ )和讲授( $Z = -4.19, P = 0.000 < 0.05$ )等行为较少。另外,在应答( $P = 0.44 > 0.05$ )、生生互动( $P = 0.77 > 0.05$ )、师生互动( $P = 0.27 > 0.05$ )等行为方面,两者没有统计学意义上的差异。

## 五、研究结论

### (一) 基于大数据的课堂教学行为特征及类型

不同教师的教学风格不同、每节课的教学内容不同,这就意味着不存在单一、普适的教学行为标准。同样,对一堂课的分析结果也不能作为评价教师教学能力的依据。但是,大规模课堂教学行为的分析却能反映总样本中教师的课堂教学行为特征、类型及共性问题。

第一,每种课堂教学行为出现的频率不等且相差较大。按照均值从大到小依次为读写、讲授、巡视、听讲、应答、举手、板书、生生互动、师生互动。读写、讲授、巡视行为占比较大,生生互动、师生互动占比较少。这表明样本中的课堂教学更偏向于单向知识传递的教学方式。

第二,各行为之间具有一定的相关性。大多数教学行为之间存在一定的正相关关系,但是巡视、读写与其他行为之间存在一定的负相关。可见,教师巡视在一定程度上能够维持课堂纪律,但是对学生其他课堂行为有一定的抑制作用。另外一种解释是,课堂教学中其他行为低,教师会增加课堂巡视,由此呈现一定的负相关。此外,读写与其他行为存在一定的负相关。学生的读写行为包括一系列低头



行为,如低头阅读、低头书写等。在课堂教学中,学生长时间的低头行为可能会对其他行为产生抑制作用。

第三,教师行为、学生行为相互交织。为了研究方便,研究者最初将课堂教学行为分为教师行为和学生行为,但是实际上,教师行为和学生行为交织在一起。将课堂教学行为划分为教师行为和学生行为,忽略了教学行为之间的关联性和复杂性。后续研究可以基于人工智能引擎自动标注的课堂教学行为数据集,探索教与学行为之间的复杂关系。

### (二) 课堂教学行为差异

在性别、教龄、职称等个人特征上,教师的课堂教学行为存在一定差异。具体而言,女教师的课堂有更多的举手、应答、听讲、生生互动、讲授、师生互动等教学行为;男教师的课堂上,读写和巡视行为的频率相对较高。低教龄组教师的课上有更多的巡视行为和学生听讲行为,高教龄组教师的课上有更高频率的课堂讲授行为;高级职称教师组的课上表现出更多的举手、讲授行为,一级职称组教师的课上表现出更多的板书、读写行为,二级职称组教师的课上呈现出更多的巡视、生生互动行为。

在教学行为类型上,举手、应答、听讲、读写、生生互动、板书、巡视、讲授和师生互动等课堂教学行为,在不同课堂类型之间均存在显著差异。然而,让人意外的是,生生互动、师生互动在讲授型与混合型、讲授型与对话型课堂之间没有显著差异。同样,对话型课堂的举手、应答等学生行为的频率没有显著高于讲授型课堂,原因可能是研究样本来源单一。本研究分析样本并非经过多次演练的示范公开课,也不是精品课,样本主要源于欠发达地区的常规课堂教学,其课堂教学方式以教师讲授为主。尽管本研究将课堂分为四类,但是样本中混合型课例占比较少(占2.56%)。在总样本中,教师讲授、学生课堂练习仍是主要的课堂形式,课堂教学缺乏有效的师生互动、生生互动。该地区可以采用协作学习、任务驱动等多种教学方式,加强课堂互动。

### (三) 课堂教学行为分析应用

发挥人工智能优势,推动人工智能和教育深度融合是我国人工智能教育的重要目标。人工智能支持的课堂教学行为分析以课堂教学为阵地,聚焦人工智能、物联网等技术与课堂教学的深度融合,以课

堂教学行为分类模型为理论框架,通过伴随式、常态化AI引擎,采集多模态学习行为,自动标注课堂教学行为,为教师的课堂教学表现、专业发展提供一定的实证式数据支撑(顾小清等,2021)。具体应用包括:

第一,为发现课堂教学行为规律提供方法,即通过课堂教学视频数据采集与存储、课堂行为建模和计算,为大规模量化课堂教学行为提供可能。随着样本量的增加,人工智能引擎将融入更多的多维关联的课堂教学行为特征,为挖掘课堂教学行为之间的规律、行为的变化模式提供更加客观和全面的证据。第二,为教师改进课堂教学提供客观评价。一线教师会通过多种途径学习,其中,从自己的教学实践中学习是重要的学习途径(约翰·D. 布兰思福特,2013)。人工智能技术能够实现课堂教学的常态化伴随式纵向数据采集,利用班级、教师和学生等行为建模技术,通过构建不同地区、不同学段、不同学科的课堂教学行为常模,为教师提供课堂教学的过程性、增值性评价报告,实现教学行为的智能分析。在教学实践中,教师可以将这些客观分析的结果与主观总结相结合,通过教学诊断,促进教学反思,进而改进课堂教学。第三,为教师研修活动提供支撑平台。区域教研员可以依托人工智能课堂教学行为分析系统,开展线上线下相结合的课例研修、直播录播课堂支持的同侪互助研修、主观评课数据与客观课堂分析数据相结合的混合式研修活动。这种实证数据支撑的教研活动的开展,能够为精准备课、评课议课及教学反思提供有力的抓手,最终促进教师信息化教学能力的提升(张妮等,2020)。

本研究存在如下局限:第一,在技术上,人工智能引擎自动标注识别的精度有待提高。课堂教学行为智能采集系统可通过标注、分析和融合教学视频中的语言、手势、姿态等多模态数据,丰富教学行为识别类型,提高教学行为识别精度;第二,在数据源上,本研究主要分析师生课堂行为类型等观察性指标,未涉及教师和学生的认知、情感等内在心理状态。未来研究可以将课堂参与主体的认知和情感数据,与课堂全景式视频相结合,更加全面立体地分析课堂教学行为规律;最后,值得注意的是,人工智能应用于教育必然会产生及使用大量的敏感数据,如何保护学生、教师的个人隐私,以及保证数据的安全

性是一个不容忽视的伦理、利益和安全问题(唐汉卫, 2018), 这也是未来需要关注的重要研究方向。

#### [注释]

①师生互动既有教师行为也有学生行为, 两种行为交织一起。鉴于大部分课堂中师生互动由教师主导发生, 为了分析方便, 本研究将师生互动归类为教师行为。

#### [参考文献]

- [1] Brown, L. (2020). What is verbal and non-verbal communication? [EB/OL]. [2020-9-11]. <https://lexiehearing.com/us/library/verbal-and-non-verbal-communication/>.
- [2] 程云(2015). 基于视频的课堂教学行为分析方法研究[D]. 武汉: 华中师范大学: 32.
- [3] 程云, 刘清堂, 王艳丽等(2017). 课堂教学行为分析云模型的构建与应用研究[J]. 远程教育杂志, 35(2): 36-42.
- [4] 陈松菁(2014). 基于 MMA Video2.0 的视频语篇分析应用研究[J]. 长春大学学报, 24(8): 1139-1143, 1147.
- [5] 方海光, 高辰柱, 陈佳(2012). 改进型弗兰德斯互动分析系统及其应用[J]. 中国电化教育, (10): 109-113.
- [6] Flanders, N. A. (1963). Intent, action and feedback: A preparation for teaching [J]. Journal of Teacher Education, 14(3): 251-260.
- [7] 傅德荣, 章慧敏, 刘清堂(2011). 教育信息处理(第二版)[M]. 北京: 北京师范大学出版社: 96-105.
- [8] 国务院(2017). 国务院关于印发新一代人工智能发展规划的通知[EB/OL]. [2020-12-01]. [http://www.gov.cn/zhengce/content/2017-07/20/content\\_5211996.htm](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm).
- [9] 郭红伟, 杨雪燕(2020). 多模态视阈下教师元话语与手势语的符际关系研究[J]. 外语教学, 41(4): 52-57.
- [10] 顾小清, 王超(2021). 打开技术创新课堂教学的新窗: 刻画 AIoT 课堂应用场景[J]. 现代远程教育研究, 33(2): 3-12.
- [11] 顾小清, 王炜(2004). 支持教师专业发展的课堂分析技术新探索[J]. 中国电化教育, (7): 18-21.
- [12] 何文涛, 黄立新, 王良辉(2019). 教学过程数据驱动下的教学解释探究[J]. 中国电化教育, (7): 38-45.
- [13] 何文涛, 杨开城, 张慧慧(2018). 智慧教室环境下协作学习的运行特征分析[J]. 中国电化教育, (8): 45-53.
- [14] 蒋立兵, 毛齐明, 万真等(2018). 智慧教室促进高校课堂教学变革的绩效研究: 基于课堂教学行为的分析[J]. 中国电化教育, (6): 52-58.
- [15] Ku, O., Liang, J. K., Chang, S. B., & Wu, M. (2018). Sokrates teaching analytics system (STAS): An automatic teaching behavior analysis system for facilitating teacher professional development [C]// Proceedings of International Conference on Computers in Education. Manila, Philippines: Asia-Pacific Society for Computers in Education: 696-705.
- [16] Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system[J]. Behavior Research Methods, 41(3): 841-849.
- [17] Lesson Study Alliance. (2020). LessonNote-A free iPad app to support lesson study [EB/OL]. [2020-12-01]. <http://www.lsalliance.org/lessonnote/>.
- [18] Lim, V. F. (2017). Analysing the teachers' use of gestures in the classroom: A systemic functional multimodal discourse analysis approach[J]. Social Semiotics, 29(1): 83-111.
- [19] Lim, V. F. (2019). Investigating intersemiosis: A systemic functional multimodal discourse analysis of the relationship between language and gesture in classroom discourse[J]. Visual Communication, 20(1): 1-25.
- [20] 刘凯, 胡祥恩, 王培(2018). 机器也需教育? 论通用人工智能与教育学的革新[J]. 开放教育研究, 24(1): 10-15.
- [21] 刘清堂, 何皓怡, 吴林静等(2019). 基于人工智能的课堂教学行为分析方法及其应用[J]. 中国电化教育, (9): 13-21.
- [22] 刘向永, 李傲雪, 付奕宁等(2018). 基于电子书包的小学英语课堂师生互动分析: 以“How are you?”单元为例[J]. 电化教育研究, 39(8): 97-102, 121.
- [23] 刘喆, 苏新冰, 杜炫杰(2020). 智慧教室环境下的数学课堂教学行为研究[J]. 数学教育学报, 29(4): 44-51.
- [24] Mintzes, J. J. (2010). Relationships between student perceptions of teaching behavior and learning outcomes in college biology [J]. Journal of Research in Science Teaching, 19(9): 789-794.
- [25] 宁虹, 武金红(2003). 建立数量结构与意义理解的联系: 弗兰德互动分析技术的改进运用[J]. 教育研究, (5): 23-27.
- [26] Prieto, L. P., Sharma K, Dillenbourg, P., & Jesús, M. (2016). Teaching analytics: towards automatic extraction of orchestration graphs using wearable sensors[C]// Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. New York: ACM: 148-157.
- [27] 任婕(2019). 基于多模态身体行为信息融合的学生学习兴趣分析方法研究[J]. 仪器仪表用户, 26(12): 109-112.
- [28] 司治国, 杨卉, 王陆(2012). 英语新手教师教学行为监控与分析的实证研究[J]. 中国电化教育, (10): 66-69.
- [29] 孙众, 吕恺悦, 骆力明等(2020). 基于人工智能的课堂教学分析[J]. 中国电化教育, (10): 15-23.
- [30] 孙众, 马玉慧(2013). 视频案例多元分析法促进新手教师专业发展研究[J]. 电化教育研究, 34(8): 104-109.
- [31] 唐汉卫(2018). 人工智能时代教育将如何存在[J]. 教育研究, 39(11): 18-24.
- [32] 王陆, 彭功(2020). 2015-2019 年中小学课堂高阶问题特征图谱[J]. 电化教育研究, 41(10): 65-72, 113.
- [33] 王陆, 张敏霞(2019). 基于课堂教学行为大数据的课堂观察方法与技术[M]. 北京: 北京师范大学出版社: 2.
- [34] 吴权威, 梁仁楷, 古骥毓等(2020). 苏格拉底教学分析与分享平台: 一个促进教师专业成长的教研平台[EB/OL]. [2020-11-10] <https://www.habook.com.cn/academic.php?act=view&id=43>.
- [35] 杨平展, 刘娟文, 罗平(2009). 关于教学行为分类系统

VICS 的讨论[J]. 教师, (6): 3-4.

[36] 约翰·D. 布兰思福特(2013). 人是如何学习的: 大脑、心理、经验及学校[M]. 上海: 华东师范大学出版社: 170-172.

[37] 余胜泉(2018). 人工智能教师的未来角色[J]. 开放教育研究, 24(1): 16-28.

[38] 郁晓华, 黄沁(2018). 学习分析视角下的数字化课堂互动优化研究[J]. 中国电化教育, (2): 12-20.

[39] 张妮, 刘清堂, 徐彪等(2020). 支持教师区域研修的 PAST 模型构建及应用研究[J]. 中国电化教育, (4): 93-101.

[40] 张屹, 祝园, 白清玉等(2016). 智慧教室环境下小学数学课堂教学互动行为特征研究[J]. 中国电化教育, (6): 43-48, 64.

[41] 赵刚, 朱文娟, 胡碧灵等(2020). 多维度特征融合的教学行为智能分析模式构建[J]. 电化教育研究, 41(10): 36-44.

[42] 周鹏霄, 邓伟, 郭培育等(2018). 课堂教学视频中的 S-T 行为智能识别研究[J]. 现代教育技术, 28(6): 54-59.

(编辑: 魏志慧)

## Automated Annotation of Classroom Behaviours with AI Engine

LU Guoqing<sup>1</sup>, XIE Kui<sup>2</sup>, LIU Qingtang<sup>1</sup>, ZHANG Chenwen<sup>3</sup> & YU Shuang<sup>1</sup>

(1. Hubei Research Center for Educational Informationization, Central China Normal University, Wuhan 430079, China; 2. The Research Laboratory for Digital Learning, The Ohio State University, Columbus, OH, USA, 43210; 3. Hubei University of Arts and Science, Xiangyang 441053, China)

**Abstract:** Classroom instructional behavior is an important factor that affects instruction and learning, while the existing instructional behavior labeling has some problems, such as labor-intensive, fuzzy classification, and complex coding. Artificial intelligence (AI) technology provides new opportunities for seamless and anytime data collection of instructional behaviors. Based on sorting out the evolution of instructional behavior labeling, this study focuses on exploring the characteristics of large-scale classroom instructional behavior and the difference of classroom instructional behavior in different teachers and instructional models. Research samples included 1201 classroom videos from three junior high schools in the northwestern region of China. An AI engine was used to automatically mark teaching and learning behaviors, and obtain a large sample of classroom behavior data. Then, correlation analysis, principal component analysis, and difference analysis methods were used to analyze the characteristics, type, and difference of classroom behaviors. The study found that: 1) The frequency of instructional behaviors varied widely. Among them, reading-writing, lecturing, and inspection behaviors accounted for a large proportion of all behaviors. On the other hand, there were relatively fewer student-student and teacher-student interactions; 2) Most behaviors related to each other. For example, the student response and student-student interaction had the highest correlation score. However, teacher inspection, reading-writing, and other behaviors were negatively correlated; 3) Two factors were extracted through principal component analysis, and teacher behaviors and student behaviors were not completely different components; 4) Teachers with different characteristics showed differences in their instructional behaviors; 5) There were differences between instructional behaviors in different instructional models. The research provides a method and practical reference for finding the characteristic of instructional behavior, improving classroom instruction, and carrying out teaching and research activities in the era of AI.

**Key words:** artificial intelligence; instructional behavior; smart annotation; correlation analysis; difference analysis