

教师评语知多少?

——探析文本后面的情感价值

郁晓华, 战晓瑜

(华东师范大学 教育信息技术学系, 上海 200062)

[摘要] 随着教育评价改革的深化,教师评语作为重要的数据分析来源,虽带有丰富的情感色彩,但其非结构化的特点导致分析并不深入。采用细粒度的文本情感分析方法,基于教师属性和学生属性的情感识别,90237条教师评语文本背后的新价值被重新加以探析。研究构建了教师风格、教师期望和学生学习力三大情感词典,定义了六种教师风格和四个等级的教师期望,再设计情感公式计算了学生的学习力得分,并基于该得分将学生聚为六个类群。然后,结合学生的学习成绩探讨了不同教师风格、不同教师期望和不同学习力水平之间的师生匹配策略。研究发现相较于低学习力水平的学生而言,学习力水平高的学生更适合由温和且民主的教师指导,同时,教师评语中的低期望信息对他们学习产生的负面影响也相对较小。研究成果可引入相关教学支持系统,协助利用教师评语进行师生画像,还可为师生匹配决策提供支持。

[关键词] 学习分析;教师评语;文本情感分析;教师期望;教师风格;学习力

[中图分类号] G434 **[文献标志码]** A

[作者简介] 郁晓华(1977—),女,浙江萧山人。副教授,博士,主要从事信息化教与学、学习技术系统设计与开发、学习分析与可视化、计算思维培养与测评等研究。E-mail:xhyu@deit.ecnu.edu.cn。

一、引言

随着2020年《深化新时代教育评价改革总体方案》等一系列政策的出台^[1],学习分析成为教育评价的创新支点。教师评价以教师评语为主要载体,是教师将学生当前的学习表现与其教育目标进行比较后所提供的反馈^[2],有时还会包含学生接下来要如何改进的建议类信息。它是反映学生学习过程和成效的一类特殊的师生交互数据,是教育评价分析的重要数据来源。在智能教育时代,教师应用各种平台工具开展教学评价,积累了大量的评语数据。这些评语数据有着丰富的内涵,但相关分析研究开展的程度并不高,更没深入文本挖掘评语中丰富的情感属性。如果对教师评语加以充分分析,可从中获得教师和多维度的

信息。比如,可以从评语的写作模式、语气和用词等角度解读教师的风格^[3],了解教师对学生的期望^[4]。除此之外,教师评语描绘了学生在注意力、投入度等方面的学习表现^[5],通过对这些信息的挖掘还可揭示学生的学习力水平。

但教师评语文本的非结构性特点和丰富语义是当前阻碍解读的最主要原因。各学科教师在评价学生时会结合学科特点和自身的写作风格反映学生的不同学习问题,这些评语结构复杂多变,风格迥异,很难进行常规性数据统计和分析。本文采用一种细粒度的文本情感分析方法,基于属性的情感识别,展示了教师评语背后的新价值,包括教师风格的揭示、对学生学习力的画像、师生匹配策略探索等。这些信息,一方面可以帮助教师认识到评语写作对学生学习的重要

影响,提升评语写作质量;另一方面可用于指导教师评语的自动化生成,减轻教师写作评语的负担。此外,用教师评语协助对教师风格和学生的学习力进行画像,还可为当前各类教学平台中增强师生的有效互动以及促进学生的有效学习提供决策支持。

二、教师评语及其隐含的信息

(一)教师评语的概念及意义

教师评语是教师对学生某一阶段发展状态的较为全面且富有个性的质性评价。教师基于自己对学生成长过程的观察和记录,对学生的发展变化作出评价^[6]。教师评语不仅描述了学生的学习任务完成情况,还会对学生的态度和行为表现加以评价,有些评语还会包含对学生接下来要如何改进的建议类信息。Walberg 在对影响学生学业表现的 26 项教师因素进行研究后,发现教师评语位列第三^[7]。教师评语能够引导学生的学习行为,激发其内在发展的动力,还能够帮助学生客观地认识自我、建立自信,形成自我概念^[8]。教师评语也是教师向学生传达情感的一种重要渠道,对学生的情感发展至关重要。Noels 等指出教师的话语风格会影响学生的内在学习动机^[9],Stevick 则通过研究发现评语引发的情感反应将影响学生长时记忆的形成及重构^[10]。

教师评语一般以描述性文字居多,有时也会配一些符号元素,比如笑脸。本研究中的教师评语属于前者,是教师输出的对学生当天或一周内学习表现的形成性反馈和进一步学习的建议。从评语的情感属性来看,可能是对学生学习表现的表扬和对学生个人学习能力的赞赏,也可能是某种程度的否定和批评。

(二)教师评语中隐含的信息

1. 评语视角下的教师风格

教师风格是教师个人结构中的稳定构成因素,其形成受到教师个人性格特点、教师职业技能储备及教学经验的影响^[11]。Grigorenko 等将教学风格归结为七种类型,分别是立法型、执行型、评判型、整体型、局部型、激进型和保守型^[12]。不同教师风格能在不同程度上吸引不同类型的学生,让他们积极地参与课堂活动,激发他们的创造力。

教师在撰写评语时往往会带有情感色彩,体现其教师风格。Nurmukhamedov 等发现,在评语中擅长使用模糊限制语(如情态动词“可能”)的教师往往比较委婉含蓄,因为这类词语能够缓和批判性评语带来的人际关系紧张^[13]。相反,某些情态动词的使用则表现了教师的权威性和支配力^[14],比如“必须”“一定”

等。戴炜华等提出我们可以从教师对学生的称谓中看出教师所持的视角是以教师为中心的还是以学生为中心的^[15]。如果对学生的称呼多为“你”“该同学”等,学生实际上是处于一种“被评价”的地位,他们的积极性也很难被调动。除此之外,教师评语所包含的信息容量也体现了教师风格中的细致程度。有的教师比较注重细节和内容的深度,会从作业、课堂互动、社会交往等方方面面对学生展开评述,有的教师则喜欢采用全局性和抽象性的视角,仅仅用一两句话就概括了学生的近期表现。

2. 教师评语中的教师期望

教师根据教学经验,会对学生的学习特征和行为进行表征和画像,并进一步内化为一种心理图式,影响其针对不同成就水平学生的教学方法的选取。高成就水平的学生往往被视为对学习感兴趣的、积极性很高的、努力的、守纪律的,教师会将他们与积极的情感相关联。相反,低成就水平的学生就被看作是不遵守纪律的、对学习没有热情的、不聪明的,教师会将他们与消极的情感相关联。然后,教师会依据自己对高、低成就水平学生的这种心理表征来预测学生未来获得的成就,构建不同的教师期望。

罗森塔尔在 1973 年提出了一个教师期望四因素模型,该模型认为教师期望会从学习内容、学习机会、评语和互动氛围四个方面影响教师行为^[16]。在评语方面,罗森塔尔认为教师在向高期望学生输出评语时,会聚焦在学生的学习成效上,而不太关注课堂中的不良行为表现,评语内容会更加细致且具有建设性,评语的整体基调也会更加积极;反之则刚好相反。

3. 教师评语中的学生信息

由于教师评语的根本目的是反映和诊断学生的学习情况,因此评价的主体始终是学生。评语中含有丰富的学生信息,除了课程知识的掌握情况外,还包括学生的学习能力、学习兴趣、学习积极性以及在日常学习行为中表现出来的自制力和自控性等^[17]。教师自身教学经验的不同,或者教师关注点的不同,都会影响教师评语中所反映学生问题的全面程度和细致水平。另外,学生在不同时期、不同学科中的学习表现也存在差异。只有将这些不同来源的评语所揭示的学生信息加以汇总,才可对学生的学习表现形成一个较为全面的勾画。

三、教师评语的现有研究内容与分析方法

(一)教师评语的现有研究内容

现有的教师评语研究,一般的做法都是先按照一

定的标准对评语进行分类,然后再比较不同类型教师评语对学生学习的促进效果。评语分析的主要分类依据有评语的信息量、及时性、效价、关注点和功能等。Dempsey 等按照评语中包含的信息量对评语进行了等级划分,从无内容评语到简单的对错反馈,到正确答案的告知,再到对正确(错误)答案的解释^[18]。王林海等从功能的角度将教师评语分为指出问题型和提出建议型,并发现疑问句提出建议型评语对学生作出实质性改正更加有效^[19]。Kluger 等指出不同效价的教师评语对学习成绩的影响不同,正向的评语能够激励学生坚持完成任务,负面的评语则会削弱学生的学习毅力^[20]。Hattie 等对 74 项教师评语研究进行元分析发现,最有效的评语形式是那些提供改进线索或强化策略给学习者的评语,以及与学习目标相关的评语,与教学过程相关、纯粹的表扬或批评的教师评语对提高成就的效果最差^[21]。这些研究虽然比较了不同类型评语对学习的影响差异,但并未触及评语所传递的教师情感信息,以及评语影响效果中的师生匹配问题。

也有不少研究开始摆脱简单的类型定位分析范畴,从评语的形式、内容丰富度和语言用词等角度对教师评语的构成及其作用机制展开了探究。比如,刘祎通过对教师评语中关键词词频的统计归纳了教师评价的关注点^[22];杨维倩通过对教师评语中指标(学习方面、智力方面、个性方面等)的使用情况进行统计,探究教师评语的区分度^[23];孙瑜对教师评语的用词进行分析,将教师评语分为含蓄型、生动形象型和善型和指向具体型^[24]。其实,通过分析教师话语构成的内在属性,可以将评语的话语特点与更深层次的教师风格和期望相关联,还可借此探讨其中的师生权力关系^[25]。除此之外,就像学生评教文本可以为教师画像^[26],教师评语也可用于学生画像,但在此方面的相关研究却非常少。

(二)教师评语的现有分析方法

当前,教师评语研究主要用到的方法有问卷调查法、访谈法和文本分析法三种。前两者并不直接研究评语文本自身,而是通过评语撰写者的反馈来获得与评语相关的信息,比如陈朝晖采用问卷调查和访谈相结合的方式对评语的写作内容、结果的呈现和利用等进行了解^[27]。虽然文本分析法可用于解读评语文本自身,比如孙瑜对评语中教师的评价价值观和评价理念进行了分析^[24],但却无法解读出教师评语中丰富的情感色彩。因此在本研究中,我们采用了文本情感分析法。文本情感分析也称为观点挖掘和观点提取,是对带有情感色彩的主观文本的分析、处理、归纳和推理

的过程^[28]。文本情感分析主要分为情绪的极化分析计算与观点倾向性分析计算两个方面^[29],实现的方法主要分为基于机器学习的方法和基于词典的方法。

情感分析法在教育领域的应用十分丰富,主要用于教学设计的改进、问题学生的监督和预测、智能教学系统的构建等方面。比如,Altrabsheh 等从学生在课堂上的实时评论中提取情感词,然后计算消极词汇(如困惑、复杂、没听懂)所占的比例,以此来调整教师的上课节奏^[30]。Wen 等从 MOOC 讨论区的帖子中挖掘学生对课程和工具的看法,以此预测学生的辍学率^[31]。Ortigosa 等构建了具备情感感知的智能导师系统,向有消极情绪的学生推荐个性化活动来激发他们的学习动机^[32]。但这些情感分析研究的数据大多为学生在学习平台或教学系统中发布的文本。如今,随着移动智能终端在常规教学中的普及,以及国家深化教育评价改革的推进,教师针对学生当天(周)的学习表现进行评价操作也越发便捷、灵活。这些文本也应当是情感分析的重要数据来源,加以挖掘可形成与学生话语不同的应用思路。

四、文本情感分析法解读教师评语的实现过程

(一)基于文本情感分析法的教师评语解读过程

教师评语文本情感分析主要包含以下四个步骤,情感信息的抽取、情感信息的归类、情感词典的建立和情感信息的归纳。情感信息的抽取是对教师评语语句中情感词的提取;情感信息的分类指根据语义识别情感词所描述的学生属性或教师属性;将情感词与属性词进行匹配,可建立不同情感词典;情感信息的归纳指的是借助情感计算的方法对主观文本中的多个情感词进行综合考量,从而判断整句评语中各个观点的情感极性和强弱。

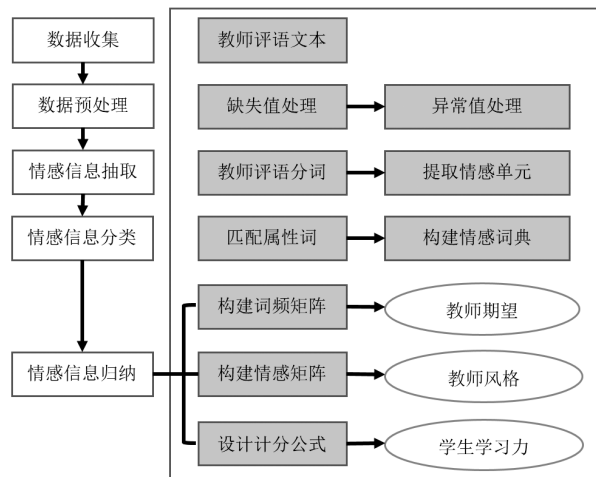


图1 教师评语情感分析路径

图1展示了从数据收集到情感信息归纳的全过程。首先收集教师评语文本,对其中的异常评语和空评语进行预处理;然后,选取合适的分词工具提取教师评语文本中的情感词汇单元;再根据情感词的语义匹配属性词,建立教师评语情感词典;最后,借助情感计算,如矩阵构建、聚类的方法对情感信息进行归纳和总结。

(二)教师评语的数据来源与分词

本研究的数据来源为上海市某一对一教育机构一个学期的教师对学生的评语数据,共包含了1226位教师对10559名学生的90237条评语,涉及语文、数学、英语、物理、化学、地理、历史、生物8个学科。

分词是文本情感分析的第一步。分词前,研究随机多次采样用人工分词与机器分词比对的方法对Jieba的原始词典进行了调整和优化,以避免分词粒度过细,并补充了一些常用词汇,以提升分词的效率。分词后共产生7万多词汇,其中,词频数大于10的词语有15624个。

(三)不同视角下情感词典的构建

由于教师在评语中会使用大量描述学生学习状况的专有名词,而目前尚未建立教师评语专用词典,因此研究挑选了词频大于10的词汇,从教师风格、教师期望和学生的学习表现三个视角构建情感词典。

教师风格在评语上的体现可从教师对学生的称谓,对学生提出建议或指出问题的方式,以及教师评语自身内容的细致程度上加以考察。其中,教师提出建议的方式会体现在评语中使用情态动词的不同上。教师指出问题的方式可通过用词的委婉程度来了解。评语内容的细致程度则可通过教师评语中谈及的学生学习力相关情感词的有无来判断。由此,我们构建的教师风格词典主要依据是评语中的称谓词、指出问题用词、提出建议用词和学习力相关词等,见表1。

表1 教师风格情感词典(部分)

维度	数量	情感词汇举例
指出问题(委婉温和词)	43	有欠缺、有待加强……
指出问题(直接严厉词)	10	太差、严厉批评……
称谓词(平等)	6	名字、昵称(如小家伙)
称谓词(权威)	2	学生、该生
提建议(权威情态词)	6	必须、严格按照……
提建议(征询情态词)	17	或许、希望……
学习力词(细致)	780	自信心强、忘性大…… (参考学习力情感词典)

教师期望可以从教师评语关注的学习方面、教师评语的极性以及建设性意见的有无来体现。教师在向

低期望学生输出评语时,会更更多地关注学生在课堂上的不良行为表现,评语的极性也会偏消极,一般不会附以改进建议;而面对高期望学生时,教师会更更多地描述学生的学习成效,评语的极性会更加积极,内容也会更加丰富且具有建设性。因此教师期望词典主要从成效词、行为词、表扬词、批评词和建议词五个维度构建,见表2。

表2 教师期望情感词典(部分)

维度	数量	情感词汇举例
行为词	327	东张西望、坐不住……
成效词	453	准确无误、超额完成……
表扬词	513	情绪高涨、反应快……(学习力情感词典中的所有正向词)
批评词	267	听不进去、缺乏自信……(学习力情感词典中的所有负向词)
建议词	122	多笔算、加紧练习……(具体的建议词)

表3 学生学习力情感词典(部分)

一级维度	二级维度	数量	效价	情感词汇举例
学习动力	情绪智力	70	正	情绪稳定、情绪饱满……
			负	不愿听课、闹情绪……
	学习兴趣	43	正	兴趣浓厚、很爱学……
			负	厌学情绪、不感兴趣……
学习能力	理解表达	84	正	能审清、反应快……
			负	啰嗦、词不达意……
	策略反思	81	正	灵活运用、举一反三、学以致用……
			负	死记硬背、生搬硬套、死算……
学习毅力	自觉性	103	正	按时完成、提前完成……
			负	潦草、打瞌睡……
	自控性	89	正	静下心来、守纪律……
			负	东张西望、静不下来……

在构建学生学习表现情感词典时,研究借鉴了学习力的概念。学习力描述了学生获取知识、分享知识、运用知识和创造知识的能力,是一个人的学习动力、学习毅力和学习能力的总和^[33-34]。学习动力主要指有明确的目标、信心和想法,这些是学生获得知识和提高能力的先决条件;学习能力是指学生能独立学习、解决问题,以及获取、整合和探索新知识的能力;学习毅力描述了学生在学习过程中的耐力和意志。学习动力维度下又可细分为情绪智力、成就动机、学习兴趣、学业效能四个二级维度^[35];学习能力维度可细分为归纳推理、理解表达、记忆能力、注意力、计划性、策略反

思、创新批判七个二级维度^[36];学习毅力维度可细分为自觉性、自控性、目标性和坚韧性四个二级维度。学习力情感词典示例见表3。

最后,研究采用德尔菲法,经过四轮专家反馈和修改,最终形成了教师风格、教师期望和学习力三个情感词典。有一部分词会重复出现在多个词典中,是因为基于不同视角的理解,这些词可以反映不同的教师属性或学生属性。

五、文本情感分析下教师评语价值的新发现

本研究对教师评语新价值展开的探索如图2所示。对于教师风格和教师期望情感词典中的情感词,首先索引出该情感词描述的属性词,随后建立情感矩阵和词频矩阵标记矩阵值,最后返回教师风格和教师期望。对于学生学习力情感词典中的情感词,首先对情感词汇进行极性和权重的标注,然后基于权重值和极性值计算学生的学习力各维度情感得分,接着再通过聚类的方式归纳出不同学习表现的群体特征。最后,再结合学生的学习成绩,探讨教师期望、教师风格和学习力表现之间的匹配策略。

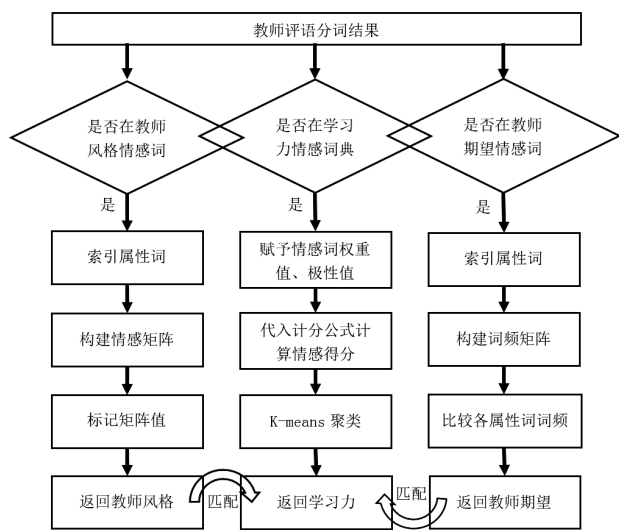


图2 教师评语情感分析应用路径

(一)教师风格标记

在教师风格情感词典中,教师指出学生问题的方式有委婉和直接之分,分别体现的是温和型和严厉型的教师风格;教师对学生的称谓以及教师对学生提出建议的方式有平等和权威之分,分别体现的是民主型和权威型的教师风格;而教师对学生的评价有一言以概之的,还有从各个学习方面具体而论的,则分别体现的是粗泛型和细致型的教师风格。这样,共得到六种类型教师风格。

在对教师风格进行标记时,主要采用构建情感矩

阵的方式^[37]。由于教师风格的判定与情感词的词频无关,与情感词的有无直接相关,因此当教师评语中出现了某一风格维度下的情感词时,则返回相应的教师风格类型。如果一位教师的评语中同时出现了表征两种风格相对的情感词汇时,则不作标记。本研究共对1226位教师的风格进行了标记,其中有342位为“粗泛型”,60位为“严厉型”,540位为“权威型”,161位为“温和型”,884位为“细致型”,429位为“民主型”。还有11位被同时标记为“细致+权威+严厉”型,257位被同时标记为“细致+权威”型。

为了验证情感分析结果的准确性,研究从六类不同风格的教师中随机抽取了共计76位教师,要求他们对各自的风格符合程度进行自评(5分为满分)。除了“权威型”教师(均分为1.82)和“粗泛型”教师(均分为2.23)认为结果不太符合实际情况外,其他风格类型的教师均认为基本符合实际情况(均分在4以上)。分析原因,一是可能本研究构建的情感词典不够完整,导致部分情感词汇未被识别;二是可能教师自评结果不真实,应当通过学生评教或教师互评对教师风格进行进一步的检验。

(二)教师期望等级

在对教师期望等级进行标记时,采用的也是构建情感矩阵的方式,只是教师期望的判定与情感词的词频直接相关,因此构建的矩阵是词频矩阵。首先,将教师评语分词结果按学生ID、教师ID进行分组;随后,统计每位教师给每个学生的评语中五种教师期望属性词的词频,建立词频矩阵;最后,依据教师评语中高期望信息的多少来标记教师对学生的期望等级。本研究设定了三个条件:如果教师评语的分词结果满足所有三个条件,则该教师的期望等级标记为“高期望”;如果满足任意两个条件,则标记为“较高期望”;如果仅满足任意一个条件,则标记为“较低期望”;如果一个条件都不满足,则为“低期望”。

条件1:成效词的个数>行为词的个数

条件2:表扬词的个数>批评词的个数

条件3:建议词的个数>0

将标记结果加以统计,有9.8%的学生收到的教师评语属于高期望评语,30.2%收到的是较高期望评语,32.9%收到的是较低期望评语,27.1%收到的是低期望评语。

(三)学生学习力画像

1. 设计计分公式量化学习表现

由于情感词所表达的情感在极性和强度上存在差异,我们可以通过赋予情感词极性值和权重值的方

式来设计情感计算公式^[38],然后将各情感词的情感值相加,从而得到整句话的情感倾向得分^[39]。

为此,研究将情感词定义为褒义和贬义两种极性,再将其情感强度分为三档,由此情感词的权重值范围就为-3到+3。基于这一设定,学习力情感词典中的每个情感词的权重值和极性值在被初步标定后,经由专家进行审核和修改直至最后确定。再由于一名学生会收到来自多个学科教师的多条评语,为了综合考虑所有教师评语中对学生学习表现各维度的评价,我们设计了以下计分方法。

步骤1:累加每一条教师评语中学生在某一维度下所有积极词性的情感词的权重值,得到 S_p ;累加每一条教师评语中学生在某一维度下所有消极词性的情感词的权重值,得到 S_n 。

步骤2: S_p 和 S_n 的差值 S 就是学生在该条评语中某一维度下的总情感得分。

步骤3:学生在所有评语中某一维度下的总情感得分 S_c 就是所有评语情感得分求和后的平均值。

经过上述计分过程,所有学科教师对学生的表现各维度的评价都被考虑在内,将每位学生在二级维度得分求平均后就得到一级维度的最终得分。

2. k-means 聚类

计算出每个学生的学习力情感得分后,采用k-means聚类,学生被划分为六个类群,见表4,依次被命名为脆弱型、坚韧型、能学型、学困型、爱学型和厌学型。其中,脆弱型学生与坚韧型学生在学习毅力方面表现出两极分化,能学型学生和学困型学生则在学习能力上存在较大差异,而爱学型学生和厌学型学生在学习动力方面表现出巨大差异。

表4 各学生类群的中心值和人数分布情况

类别	学习动力标准 化情感得分	学习能力标准 化情感得分	学习毅力标准 化情感得分	人数
脆弱型	0.12	-0.09	-0.84	217
坚韧型	0.10	0.16	0.38	681
能学型	0.07	0.20	-0.03	606
学困型	0.04	-0.52	-0.09	394
爱学型	1.59	-0.00	-0.08	95
厌学型	-1.54	-0.05	-0.07	195

研究从六类学生群体中随机抽取了52名学生,将他们各学习力维度的情感得分反馈给教师进行符合程度的评价(5分为满分)。教师们的平均打分为3.81,认为学习力情感得分基本符合学生的实际学习表现。

(四)师生匹配讨论

1. 教师期望对学生学习的影响

教师对学生的不同期望,首先会影响到教师教学方法的选择;其次,学生对教师评语中教师期望的敏感度也不尽相同。敏感度高的学生对教师评语中的期望信息解读会更加准确,同时也更加容易受到教师期望的影响。为了探究教师期望对不同类型学生学习的影 响,研究对比了各类别学生在不同期望水平的教师评语作用下的成绩进步情况,见表5。结果表明,不同类型学生对教师评语中教师期望的敏感度的确存在差异。教师评语中的低期望信息不会对能学型和坚韧型学生的学习产生负面影响,反而会激励他们取得更高的成就。而脆弱型、厌学型和学困型的学生则需要高期望的评语,低期望评语会对他们的学习产生负面影响。这与Madon等^[40]和Brattesani等^[41]所作的研究结论一致,成绩中下游以及没有清晰的自我概念的学生往往受教师评语中教师期望的影响更大。

表5 不同教师期望水平下的学生成绩进步情况对比

学生类别	不同教师期望水平下的学习成绩进步情况对比
脆弱型	高期望>较低期望>较高期望>低期望
坚韧型	较低期望>较高期望>高期望>低期望
厌学型	高期望>较低期望,较高期望>低期望
爱学型	高期望,较高期望>低期望>较低期望
学困型	高期望>较高期望>较低期望>低期望
能学型	较高期望,较低期望>高期望>低期望

2. 教师风格对学生学习的影响

教师风格对不同类型学生有不同的吸引力,这种吸引力上的差异会造成学生在课堂中投入度的不同,以及对于课程学习兴趣的差异^[42]。因此,探讨不同教师风格下学生的学业成绩进步情况可用于指导师生匹配,最大程度地激发学生的学习热情和积极性,同时还可协助教师进行自身风格调整。

本研究对每种风格的教师教授的各学生类别的成绩进步率情况进行了统计,见表6。结果显示,粗泛型教师对于学困型学生和厌学型学生,权威型教师对于厌学型学生,细致型教师对于厌学型学生,严厉型教师对于坚韧型学生和学困型学生,民主型教师对于能学型学生,温和型教师对于爱学型学生的学习取得的成绩进步率相对较高。而粗泛型教师指导坚韧型学生,权威型教师指导学困型学生,细致型教师指导能学型学生,民主型教师指导脆弱型学生和学困型学生的学习时,则可能存在一定的退步。基于这些发现,我们认为学习力水平高的学生更适合由民主且温和的

教师指导,学习力水平低的学生更适合由粗泛且严厉的教师指导。

表6 不同教师风格指导下学生成绩进步情况对比

教师类别	学生类别	成绩退步	成绩保持	成绩进步	匹配建议
粗泛型	学困型	0	74.75%	25.25%	建议
	厌学型	0	74.47%	25.53%	建议
	坚韧型	1.50%	79.70%	18.80%	酌情考虑
权威型	厌学型	0	67.21%	32.79%	建议
	学困型	1.10%	80.22%	18.68%	酌情考虑
细致型	厌学型	0.89%	73.45%	25.66%	酌情考虑
	能学型	1.07%	80.34%	18.59%	酌情考虑
严厉型	坚韧型	0	63.64%	36.36%	建议
	学困型	0	69.23%	30.77%	建议
民主型	脆弱型	1.89%	98.11%	0	不建议
	能学型	0	66.96%	33.04%	建议
	学困型	2.30%	93.10%	4.60%	不建议
温和型	爱学型	0	50.00%	50.00%	建议

六、结论与展望

教师评语在教育评价中一直扮演着重要的角色,但面对教育评价改革的深化,作为重要的反映学生学习过程和成效的数据分析来源,教师评语却得不到充分的挖掘和利用。为此,本研究采用了一种细粒度的文本情感分析方法,尝试从不同视角展示了教师评语背后的新价值。研究建构了三大情感词典,从教师评

语中解读出了六种教师风格、四个等级的教师期望以及六种类型的学生学习力,最后再结合学生的学习成绩讨论了不同教师风格、不同教师期望和不同学习力水平之间的师生匹配差异。一方面,研究结论很好论证了教师评语在传递情感中对学生学习的重要影响。教师应意识到评语的重要性且正确感知学生的需求,恰当地输出对学生发展有利的个性化评语。另一方面,研究路径可作为今后教师评语研究的参考。基于属性的情感识别提炼情感词构建情感词典,再借助权重值和极性值的标注以及词频矩阵构建等方法,可将教师评语从非结构化的文本转换成量化的情感指标值,从而拓展了教师评语的分析空间。最后,研究成果可引入相关教学支持系统,以协助利用教师评语对教师风格和学生学习力进行画像,还可为师生匹配提供决策依据,以保证师生的有效互动。

本研究也有许多需要发展的地方。教师评语词典是教师评语分析的根本,但现在并没有此方面的专用词典。我们虽然采用手动标注的方式提取教师评语文本中的情感词,初步建立了教师风格、教师期望和学生学习力三个情感词典,但还存在较多情感词未被纳入词典的情况。在后续的研究中,一方面需要加大教师评语研究数据的数量和类型,另一方面还可考虑引入机器学习的方法增强词典的建设力,从而提升情感分析的准确率。未来还可基于这些情感词典,开发教师评语的自动生成系统,基于学生的不同学习表现,智能输出适合的教师评语,减轻教师写作评语的负担。

[参考文献]

- [1] 新华社.中共中央国务院关于深化新时代教育评价改革总体方案的通知[EB/OL].[2020-10-13]. http://www.gov.cn/zhengce/2020-10/13/content_5551032.htm.
- [2] ALEXANDER P A, SCHALLERT D L, HARE V C. Coming to terms: how researchers in learning and literacy talk about knowledge [J]. Review of educational research, 1991, 61(3): 315-343.
- [3] 黄焕,刘清堂,朱晓亮,王胜明,高桂平.不同教学风格的课堂话语特征分析及应用研究[J].现代教育技术,2013,23(2):27-30,36.
- [4] HARRIS M J, ROSENTHAL R. Mediation of interpersonal expectancy effects: 31 meta-analyses[J]. Psychological bulletin, 1985, 97(3): 363.
- [5] 沙丽华,崔建京,兰宁,薛达志.综合素质评价中教师评语的特点、作用及写作要求[J].辽宁教育,2019(4):43-45.
- [6] 陈惠.个性化评语的作用及撰写方法——以北京市普通高中学生综合素质评价中的教师评语为例[J].教育测量与评价:理论版,2011(10):32-35.
- [7] WALBERG H J. Improving the productivity of America's schools[J]. Educational leadership, 1984, 41(8): 19-27.
- [8] ANNETT J. Feedback and human behaviour: the effects of knowledge of results, incentives and reinforcement on learning and performance[M]. Penguin Books, 1969.
- [9] NOELS K A, CLÉMENT R, PELLETIER L G. Perceptions of teachers' communicative style and students' intrinsic and extrinsic motivation[J]. The modern language journal, 1999, 83(1): 23-34.

- [10] STEVICK E W. Affect in learning and memory: from alchemy to chemistry[J]. Affect in language learning, 1999:43-57.
- [11] 翟雪君. 教师教学风格差异的本质、结构与类型[J]. 大众文艺, 2019(19):161.
- [12] GRIGORENKO E L, STERNBERG R J. Styles of thinking, abilities, and academic performance[J]. Exceptional children, 1997, 63(3): 295-312.
- [13] NURMUKHAMEDOV U, KIM S H. 'Would you perhaps consider...': hedged comments in ESL writing [J]. ELT journal, 2010, 64(3): 272-282.
- [14] TREGLIA M O. Feedback on feedback: exploring student responses to teachers' written commentary [J]. Journal of basic writing, 2008, 27(1):105-137.
- [15] 戴炜华, 陈宇昀. 批评语篇分析的理论和方法[J]. 外语研究, 2004(4):13.
- [16] JUSSIM L, ROBUSTELLI S L, CAIN T R. Teacher expectations and self-fulfilling prophecies[M]//Handbook of motivation at school. Routledge, 2009:349-380.
- [17] 鄢超云, 魏婷.《3~6岁儿童学习与发展指南》中的学习品质解读[J]. 幼儿教育, 2013(18):1-5.
- [18] DEMPSEY J V, DRISCOLL M P, SWINDELL L K. Text-based feedback [M]// DEMPSEY J V. Interactive instruction and feedback. Educational Technology, 1993:21-54.
- [19] 王林海, 孙宁. 教师评语类型对学生作文修改类型的影响[J]. 外语界, 2007(5):38-42.
- [20] KLUGER A N, DENISI A. Feedback interventions: toward the understanding of a double-edged sword [J]. Current directions in psychological science, 1998, 7(3): 67-72.
- [21] HATTIE J, TIMPERLEY H. The power of feedback[J]. Review of educational research, 2007, 77(1): 81-112.
- [22] 刘伟. 基于学生综合素质发展的小学操行评语研究[D]. 上海: 华东师范大学, 2019.
- [23] 杨维倩. 中学班主任评语的分析与研究[D]. 大连: 辽宁师范大学, 2008.
- [24] 孙瑜. 小学生操行评语文本研究[D]. 成都: 四川师范大学, 2015.
- [25] 郭黛宝. 教师评语中的师生互动研究[D]. 北京: 首都师范大学, 2014.
- [26] 陈尧. 教师画像与评分系统的设计与实现[D]. 重庆: 重庆大学, 2018.
- [27] 陈朝晖. 普通高中学生综合素质评价实施研究[D]. 开封: 河南大学, 2016.
- [28] 刘毓, 赵云阁. 基于教学评价的中文短文情感分析[J]. 现代电子技术, 2019, 42(6):30-33, 37.
- [29] 饶元, 吴连伟, 王一鸣, 冯聪. 基于语义分析的情感计算技术研究进展[J]. 软件学报, 2018, 29(8):2397-2426.
- [30] ALTRABSHEH N, COCEA M, FALLAHKHAIR S. Learning sentiment from students' feedback for real-time interventions in classrooms[C]//International conference on adaptive and intelligent systems. Springer, Cham, 2014: 40-49.
- [31] WEN M, YANG D, ROSE C. Sentiment Analysis in MOOC Discussion Forums: What does it tell us?[C]. Educational data mining 2014, 2014:130-137.
- [32] ORTIGOSA A, MARTÍN J M, CARRO R M. Sentiment analysis in Facebook and its application to e-learning [J]. Computers in human behavior, 2014, 31: 527-541.
- [33] 裴娣娜. 学习力: 诠释学生学习与发展的新视野[J]. 课程·教材·教法, 2016(7):3-9.
- [34] 吕晓娟. 基于学生学习力的翻转课堂教学设计[J]. 电化教育研究, 2015, 272(12):98-102.
- [35] 陈维维, 杨欢. 教育领域学习力研究的现状和发展趋势[J]. 开放教育研究, 2010, 16(2):40-46.
- [36] 胡爽, 陈秉初. 学习力“六要素”的解析与分层[J]. 教育现代化, 2015(1):55-59.
- [37] 叶佳鑫, 熊回香, 杨滋荣, 童兆莉. 关键词词频及语义特征对科技文献聚类的影响研究[J]. 情报科学, 2021, 39(8):156-163.
- [38] TABOADA M, BROOKE J, TOFILOSKI M, VOLL K, STEDE M. Lexicon-based methods for sentiment analysis [J]. Computational linguistics, 2011, 37(2): 267-307.
- [39] 陈国兰. 基于情感词典与语义规则的微博情感分析[J]. 情报探索, 2016(2):1-6.
- [40] MADON S, JUSSIM L, ECCLES J. In search of the powerful self-fulfilling prophecy[J]. Journal of personality and social psychology, 1997, 72(4): 791.
- [41] BRATTESANI K A, WEINSTEIN R S, MARSHALL H H. Student perceptions of differential teacher treatment as moderators of teacher expectation effects[J]. Journal of educational psychology, 1984, 76(2): 236.

[42] 陈文革.教师评语与师生关系的构建[J].话语研究论丛,2015(1):73-86.

How Much do We Know about Teacher Comments? —Exploring Emotional Value behind the Text

YU Xiaohua, ZHAN Xiaoyu

(Department of Education Information Technology, East China Normal University, Shanghai 200062)

[Abstract] With the deepening of educational evaluation reform, teacher comments, as an important source of data analysis, are rich in sentiment, but their unstructured characteristics leads less in-depth analysis. Using a fine-grained text sentiment analysis method, based on the sentiment recognition of teacher attributes and student attributes, the new value behind 90237 teacher comments are re-explored. First, three major affective lexicons of teacher style, teacher expectation and student learning power are constructed, and six types of teacher styles and four levels of teacher expectations are defined. Next, the affective formula is designed to calculate students' learning power scores, and students are clustered into six clusters based on this score. Finally, the teacher-student matching strategies between different teacher styles, different teacher expectations, and different levels of learning power are explored in combination with students' academic achievements. It is found that students with high levels of learning power are more suitable to be guided by moderate and democratic teachers than those with low levels of learning power. At the same time, teacher comments that convey low-expectation messages have less negative impact on their learning. The results of the study can be introduced into relevant instructional support systems to assist in teacher-student profiling using teacher comments, and can also provide support for teacher-student matching decisions.

[Keywords] Learning Analytics; Teacher Comments; Text Sentiment Analysis; Teacher Expectations; Teacher Style; Learning Power

(上接第 96 页)

"Modern Educational Technology", this paper conducts a dynamic analysis of the regulation activities, regulation patterns and development trajectories of learners with different academic achievements in self-regulated learning. The results show that (1) high-performance learners have more regulatory behaviors such as pre-planning and self-reflection, but there is no significant difference between high-performance learners and low-performance learners in the behavioral performance stage. (2) high-performance learners have more self-assessment behaviors and more in-depth self-reflection and regulation patterns. (3) High-performance learners can generate value identification and stimulate learning interest at the pre-planning stage, whereas low-performance learners' value identification is more delayed. In view of the above differences, this paper gives the corresponding pedagogical enlightenments, including providing low-performance learners with learning scaffolding in the pre-planning stage and self-reflection stage, guiding low-performance learners to realize value identification of tasks in the task initiation stage, and guiding and assisting low-performance learners to carry out self-assessment etc.

[Keywords] Self-regulated Learning; Collaborative Learning Activities; Epistemic Network Analysis; Trajectory Analysis