OBE理念下AI赋能《英语演讲与辩论》四维教学评价体系构建

肖书阳
广东科技学院 国际商务与语言学院，中国·广东 东莞 523349

摘要：英语演讲、辩论课程的学习效果体现在讲稿的打磨，现场展演，互动问答，辩论交锋等多个方面，不可用一个维度的成绩来体现学生的全过程表现。目前所用的考核准则多为语言的精确度以及现场表现，对论据合理的使用、有针对性地驳斥以及团队协作等指标则考察较少。为了改善上述情况，本次评价方案新增加了语言综合运用能力、内容架构和逻辑思辨水平、交际策略及临场应变能力和跨文化认知与文化敏感度四个观察维度，并且考核部分以及所占比例没有改变。教师根据不同的讲稿、学生的自我评价材料、演讲视频和辩论的过程记录来进行综合判断；AI工具只是把文字转换成语音，对简单的排列组合进行辅助，并不会影响最终评分的结果。目前，本文提出的设计还未应用到课堂教学，评价量表等级划分、教师复核方式和技术应用流程还需在实践中进行调整与改善。

关键词：OBE 理念；AI 赋能；英语演讲与辩论；四维评价；人机协同

Developing an AI–Assisted Four–Dimensional Assessment System for English Public Speaking and Debate within the OBE Framework

Xiao Shuyang
School of International Business and Language Studies, Guangdong University of Science and Technology, China
Guangdong Dongguan 523349

Abstract: Learning in English Public Speaking and Debate is distributed across script revision, live presentations, questioning, and debate interaction, and no single score captures all of these processes. A comparison of the syllabus, teaching plan, and continuous assessment guidelines shows that linguistic accuracy and delivery are already well documented, whereas evidence use, rebuttal, and collaboration are described in less detail. The proposed framework therefore adds observations in four areas—integrated language use, content development and logical reasoning, communicative strategies and real-time performance, and intercultural awareness and sensitivity—without changing the existing assessment items or weightings. Script versions, self-evaluation reports, recorded presentations, and debate records support teachers' judgments. AI is used for transcription, material alignment, and preliminary prompts rather than direct grading. The framework has not yet been piloted; rubric boundaries, teacher review, and technical procedures require refinement in practice.

Keywords: Outcomes-based education (OBE); Artificial intelligence (AI); English public speaking and debate; Four-dimensional assessment; Human–AI collaboration

1 问题提出

《深化新时代教育评价改革总体方案》要求改进结果评价、强化过程评价，并探索增值评价、健全综合评价[1]。教育数字化相关文件也为人工智能参与评价提供了政策背景，同时强调安全有序[2]。落到语言实践课，关键不是增加自动评分工具，而是让记录说明学生学会了什么、哪些目标仍缺少证据。

《英语演讲与辩论》为英语专业必修课，共 32 学时、2 学分，课堂教学与课内实践各 16 学时。学生需完成 2—5 分钟英语演讲，并辨析、回应他人观点。定题演讲、即兴演讲和英语辩论分别安排 5、5、6 学时。授课计划将学习内容安排为十个演讲教学单元和四次辩论活动，由选题、论据和结构逐步过渡到语言呈现与辩论实践。学习成果分布在接连发生的任务中，一次展示后的总分难以完整说明。

课程设置有八个平时考核模块，贯穿整个教学过程，在将课堂表现所占比例设置为 25% 的基础上，分别设置了

演讲稿、自评报告、模仿朗读、课前演讲各占 10%，定题演讲、即兴问答、辩题简述分别为 20%、14% 和 6%。语法正确性、语句规范性和现场表现效果等可以进行量化评分，但学生选取论据、展开逻辑推理、开展辩论、团队合作等能力主要靠教师主观模糊评判。

本评价设计依靠多重理论作为支撑，学习成果是课程设计和持续改进的主要出发点[3-4]，评价任务应与课程目标相一致[5]，评价反馈要融入到学生整个学习过程[6-7]。高等教育领域有关人工智能的研究也认为，技术对学习材料的加工分析不可代替教师的专业判断[8]。上述理论原则并没有给本课程的两个实操问题提供具体的解决办法，即如何把四个观测维度融入到现有的考核项目当中，以及如何对 AI 生成的辅助提示进行教师的复核校验。该方案目前无进行教学试点，下面仅探讨设计思路和检验方法。

2 四维评价的课程依据

2.1 基于课程目标确定观测内容

课程大纲中的培养目标不能作为评价依据。以“讲好中国故事”为例，其包含了内容建构和受众适应两个方面的内容，未考虑受众需求的情况下，即使语言表达很流畅，也很难达到交际的目的。语言维度主要评价表达是否清楚、准确；内容和思辨维度看重论据对于观点的支持程度；交际与临场维度记述问答互动、团队合作和非语言表现；跨文化维度考察文化预设、受众剖析及临场应变能力。学生在公众表达、职场适应方面的发展水平，靠多项任务综合评定。

评价材料需要与教学过程同时进行，前六个单元主要是选题确定、表达目的设定、论据收集以及篇章构造等内容的整理工作，并且保留了提纲和多次讲稿作为过程记录。四次辩论任务对立论、反驳、逻辑谬误识别以及团队合作这四个方面的表现进行了全面的考察。所有观测指标以及学生的日常练习需一一对应，在期末不再增加其他量表测查。

2.2 将评价证据融入学习过程之中

证据留存可以利用课堂角色扮演、展示、演讲录音、讲稿和自评报告等方式来进行。单场演讲除了分数之外还要保留提纲、文献出处、初稿、排练笔记、音视频和修订稿。结合作业至少三次全批全改的要求，教师可以通过讲稿、录音、自评报告进行反馈，不对所有的材料重复打分，出勤纪律不作为能力证明。

评语要针对具体的材料书写，即兴问答可用于评价原始作答情况，并对回答是否完整、论证断点、肢体辅助表达情况进行标记，评价需要去掉印象分。辩题简述需对照文本核查观点、论据与逻辑而进行，流畅度只是其中一项参考。教师需对反馈标注的文稿语句或者音视频时间节点进行确定，下次优先解决一个问题来达到评价促进学习的目的[7]。

2.3 明确 AI 辅助的边界

本方案的人工智能先做机械性的工作，把超星学习通上的录音转化成文字，对齐讲稿和音视频，并标记出可能存在的停顿、语速改变或者是结构上的空缺。评分结果为以上三组。停顿可以起到强调的作用，语速会受到话题、场合的影响，教师要听后结合任务做出判断。口音、情绪识别和单一文本指标不能直接进入等级，资料真伪、跨文化表达以及最后的成绩还要由老师来确认。

3 四维指标在现有考核中的落实

表 1 为目标、任务、观察内容与补充材料的对应关系。

表 1 课程目标、考核任务与观测证据的对应

课程目标与观察维度	成绩项目	评分时关注的内容	过程材料与 AI 用途
目标 1：交际与跨文化表达	课堂表现、课前演讲、英语辩论占据比例分别为 25%、10%、10%	合作回应、文化预设等	小组记录；AI 协助汇集片段
目标 2：语言与临场表达	定题演讲、即兴问答各占 20%、14%	演讲清晰度、结构、体态	录音录像；AI 标记语速、停顿
目标 3：内容与思辨	自评报告、辩题简述各占 5%、6%	资料可靠性、连贯性	证据卡；AI 辅助定位缺漏
目标 4：综合交际表现	演讲稿、模仿阅读各占 10%	修辞是否适应情境	讲稿版本；AI 对齐材料并起草反馈

课堂表现仍旧用课堂纪律 30 分、课堂参与 70 分来记录，但是增加了发言能否接续同伴、反馈是否有用、是否根据意见进行修改等内容，并不单纯是统计次数。定题演讲 20 分沿用五级标准，优秀和良好的评价要靠证据来支撑，即推理是否完整、能否回答问题等几个方面；中等是一个观点可以辨认出但是证据单一或者局部的推理有断开的情况出现；及格的标准是观点比较明确，但所依据的资料无法验证或没有展开论证。

即兴问答的 14 分，要看学生怎样组织回答，不能脱离语境去计算语速或者停顿。辩题简述 6 分增加对论证要素和谬误的辨认；模仿阅读 10 分用于语音语调练习，不以口音像不像母语者推断综合能力。跨文化部分参照知识、态度、技能和意识的综合理解[9]，让学生先写听众分析，

然后改写同一中国故事的不同情境版本，在自评中说明取舍，模型不能代替老师对这些表达的选择。

4 随教学进度展开的评价流程

4.1 演讲准备

前六个单元的任务卡应该包含 2-5 分钟演讲目的、听众、需要提交的材料、量规和可以使用的人工智能。选题说明、听众分析和资料来源提交在提纲之前，讲稿改动留到后面处理。学生可以用 AI 检验结构上的缺失或者主张与证据的断点，而老师则对这些材料加以审核。初稿和修订稿的不同之处，就成了内容、思想的证据。

4.2 演讲展示

第七至第十单元以及定题、即兴演讲阶段的讲稿、录音、展示视频和问答记录一起放在一个任务里。转写完之后，人工智能给出来的每一个提示都要加上讲稿原句或者视频时点。系统标注一段长停顿之后，教师要回放这段话，并停顿一定时间。即兴问答要保留题目、准备时间以及原始的回答，从而体现当时的表演状态。

4.3 英语辩论

四次辩论用英国议会制的规则，材料不要多。立场卡是主张，证据卡是依据，反驳笔记记录对方的观点以及回应的位置。人工智能找出立论、驳论和追问，老师检验复述是否正确、反驳是否针对前提或者证据、队友发言是否连贯。学生修改的时候要写好原题和材料的页码、改的方法以及理由，并且可以附上反对或者问答的内容。

4.4 反馈公开

复评仍然用同一个量表，看内容、应答和协作有无实质上的改变，而不仅仅看语言是否更漂亮。平时成绩每月至少公布一次，指标反馈一起发布。多份材料显示证据不足时，在下一轮课前补充可靠性练习；反驳需重复自己的立场，分步练习需复述对方、定位前提、回应和总结。

5 实施中需要把关的三个问题

5.1 量规校准与评分偏差

试用之前，两名老师分别对一组不同水平的演讲、问答、辩论进行评分，并且讨论出五级量规的分界。AI 建议和教师判断不一致时，要标明是转写错误、说话人误标还是量规和情境理解上的偏差。不同的口音、性别和语言背景的学生也要进行对照。凡是经常给出错误标记的自动分段器，应该被弃用。

5.2 AI 使用声明与学术诚信

除了原有的抄袭扣分外，在作业的要求里加入 AI 使用声明，写明使用哪些工具进行检索、翻译、润色、语音陪练或者反馈整理，并且和提纲、草稿以及版本记录相互验证。如果教师对学生的掌握情况存有疑问，可以通过现场提问、过程说明或者即兴复述来检验，不能以 AI 文本检测的结果为依据进行评分。评价要以学生能否核实资料、解释取舍并对表达负责为标准。

5.3 音视频数据的必要限度

只收集评价必需的声音、图像和文本。上传之前要说明用途、保存期限、查看权限和替代评价方式。教学研究中删除了所有能够辨识出个体身份的信息，并且不会将可以识别的资料上传到未经安全评估的平台上去。人脸、情绪数据与本次评价无关，也不会被采集。线下提交的等其他方式的成绩数据也要保留起来，以免学生因为选择不当造成成绩下降。

6 结语

四维设计改变的是教师对于现在考核理解的角度，并不是增加成绩项目数目。讲稿、问答和辩论材料可以成为观察证据运用、思辨和合作的入口，而人工智能只是做可追溯的基础处理工作，教师还要对评分负责。本方案尚未进行课堂试用，之后还要检验量规是否实用、教师的评价接近与否、学生的改进情况如何，此后再决定指标和技术环节的留取。

参考文献：

[1] 中共中央，国务院. 深化新时代教育评价改革总体方案[Z]. 2020.

[2] 教育部等九部门. 关于加快推进教育数字化的意见：教办〔2025〕3 号[Z]. 2025.

[3] SPADY W G. Outcomes-Based Education: Critical Issues and Answers[M]. Arlington: American Association of School Administrators, 1994.

[4] 李志义. 解析工程教育专业认证的成果导向理念[J]. 中国高等教育，2014(17): 7-10.

[5] BIGGS J. Enhancing Teaching through Constructive Alignment[J]. Higher Education, 1996, 32(3): 347-364.

[6] BLACK P, WILIAM D. Assessment and Classroom Learning[J]. Assessment in Education: Principles, Policy & Practice, 1998, 5(1): 7-74.

[7] HATTIE J, TIMPERLEY H. The Power of Feedback[J]. Review of Educational Research, 2007, 77(1): 81- 112.

[8] ZAWACKI-RICHTER O, MARÍN V I, BOND M, et al. Systematic Review of Research on Artificial Intelligence Applications in Higher Education—Where Are the

Educators?[J]. International Journal of Educational Technology in Higher Education, 2019, 16: 39.

[9] BYRAM M. Teaching and Assessing Intercultural Communicative Competence[M]. Clevedon: Multilingual Matters, 1997.

基金项目：来源：广东科技学院 2025 年校级教改项目，项目名称：OBE 理念下 AI 赋能《英语演讲与辩论》“四维”教学评价研究，编号：GKZLGC2025183。

作者简介：肖书阳（1991.11.03），女，汉族，河南省驻马店市，硕士研究生，中级，研究方向：英语教育。