·住院医师规范化培训·

循证式听评课在住院医师规范化培训师资培训中的实施路径与成效分析

秦贺　陈雯婧　李欣　叶京英
清华大学附属北京清华长庚医院耳鼻咽喉头颈外科/清华大学临床医学院，北京 102218
通信作者：李欣，Email：entlixintsinghua@sina.com

【摘要】 目的　针对传统听评课模式存在的凭主观感受随意实施、点评不具体、教学功能异化等弊端，在住院医师规范化培训师资培训中系统构建循证式听评课模式，并评估效果。**方法**　以北京清华长庚医院志愿参与研究的36名住培教师为研究对象。通过成立跨学科小组、设计结构化观察表、建立具体评价标准、创造心理安全环境的循证式听评课准备阶段，以及课前碰头会、课中观摩、课后汇报的循证式听评课实施阶段，进行为期1年的听评课实践。利用混合研究方法，通过《住培教师自我效能感评估表》和《教师自我报告教学实践评估表》比较教学实践前后得分差异，以及深度访谈、自我反思日记等方面评价循证式听评课的影响。**结果**　循证式听评课效果整体显著，表现在教师教学能力及学员感受层面。36名师资的教学自我效能总分从6.21 ± 1.18上升到7.54 ± 0.89，特别是使用结构化工具进行教学分析、给予具体的描述性反馈等条目分值上升明显。行为频率总分从2.76 ± 0.65上升到3.58 ± 0.58，特别是提出高层次问题导向、引导学生临床鉴别诊断、给予及时具体的反馈等条目分值上升明显。教师访谈和反思日志共提炼出3个主题：从“害怕敌对”到“开放包容的心态”、从“基于感觉”到“基于证据”的教学分析、从“机械被动”到“主动精心”的教学实践。另外，通过焦点小组会议与文本分析呈现共同的专业话语构成学习共同体，并创建“证据诊断—协商处方—行为改造—成效诊断”的教育闭环。**结论**　循证式听评课以证据替代经验、以支持替代评价、以闭环替代孤立，提升了住培带教教师信心，优化了带教行为，为住培基地创新师资培训新方法、培训新思路提供新的启示。

【关键词】 循证式听评课；　住院医师规范化培训；　师资培训；　教学能力；　教学评价
【中图分类号】 R-05
基金项目：2023年北京市卫健委质量提高项目（2023069）
DOI: 10.3760/cma.j.cn116021-20251201-02295

Implementation pathway and effectiveness analysis of evidence-based classroom observation and evaluation in faculty development for standardized residency training

Qin He, Chen Wenjing, Li Xin, Ye Jingying
Department of Otorhinolaryngology-Head and Neck Surgery, Beijing Tsinghua Changgung Hospital/School of Clinical Medicine, Tsinghua University, Beijing 102218, China
Corresponding author: Li Xin, Email: entlixintsinghua@sina.com

【Abstract】 Objective　To address the shortcomings of the traditional observation and evaluation model, such as arbitrary implementation based on subjective feelings, vague feedback, and distorted teaching functions, a systematic evidence-based observation and evaluation model was constructed and evaluated in the training of standardized resident instructors. **Methods**　A total of 36 volunteer resident instructors from Beijing Tsinghua Changgung Hospital were selected as the study subjects. Through the preparation phase of the evidence-based observation and evaluation model—establishing interdisciplinary teams, designing structured observation forms, setting specific evaluation criteria, and creating a psychologically safe environment—and the implementation phase—pre-class meetings, in-class observations, and post-class reports—a one-year observation and evaluation practice was conducted. Using mixed research methods, the differences in scores before and after teaching practice were compared using the "Resident Instructor

Self-Efficacy Assessment Scale" and the "Teacher Self-Reported Teaching Practice Evaluation Scale", while the impact of the evidence-based model was assessed through in-depth interviews and reflective journals. **Results** The overall effectiveness of the evidence-based observation and evaluation model was significant, manifesting in both teaching competency and trainee feedback. The total self-efficacy scores of the 36 instructors rose from 6.21 ± 1.18 to 7.54 ± 0.89, with notable improvements in items such as teaching analysis using structured tools and providing specific descriptive feedback. The behavioral frequency scores increased from 2.76 ± 0.65 to 3.58 ± 0.58, particularly in items like posing high-level problem-oriented questions, guiding clinical differential diagnoses, and delivering timely, specific feedback. Teacher interviews and reflective journals identified three key themes: shifting from "fear of confrontation" to "open and inclusive attitudes," transitioning from "emotion-based" to "evidence-based" teaching analysis, and evolving from "mechanical passivity" to "proactive and meticulous teaching practice." Additionally, focus group discussions and textual analysis revealed a shared professional discourse that fostered a learning community, establishing an educational closed loop of "evidence diagnosis—negotiated prescription—behavioral transformation—effect diagnosis." **Conclusions** By replacing experience with evidence, evaluation with support, and isolation with closed loops, the evidence-based observation and evaluation model enhanced the confidence of standardized resident instructors, optimized teaching behaviors, and provided new insights for innovating faculty training methods and approaches in residency training programs.

【**Key words**】 Evidence-based classroom observation and evaluation; Standardized residency training; Faculty development; Teaching competency; Teaching evaluation

Fund program: Quality Improvement Project of Beijing Municipal Health Commission in 2023 (2023069)

DOI: 10.3760/cma.j.cn116021-20251201-02295

规范化培训是培养医学生成为临床医师的重要步骤。在规范化培训过程中，住培带教教师具有十分重要的引领作用，住培带教师资培训是影响住培带教培训效果的一个关键因素[1-2]。传统的师资培训中，教师的听评课评价是常态，但评课最终陷入“听了不评、听了不改”的尴尬境地，原因在于：第一，听评课缺乏循证逻辑，听评课是评课者基于主观经验判断，缺乏证据支撑，违背了循证“使用最佳证据来支撑实践决定”的基本逻辑[3-4]；第二，听评课功能定位是行政管理和终结评价，背离了促进教师发展这一初衷，与“以评促教、以评促学”、教师不断改进教学实践这一目的不相符，教师因怕评价而不求改变，主动学习和反思的意愿不强[5]；第三，听评课反馈过于笼统，反馈后无跟进，起不到形成性评价提示改进方向的作用[6-7]。所以，传统的听评课很难引起教师具体的改进行为。因此，本研究拟就循证式听评课在住培师资培训中的应用路径和综合应用效果展开系统性的研究，首先探究并构建循证式听评课在住培师资培训中的应用路径，即课前准备、课中观察、课后反馈和持续跟进；随后在应用实践中，从师资教学、师资自我效能感和学员学习体验等角度系统评估循证式听评课的应用效果和挑战，为住培基地创新师资培训理念和模式，提高住培质量提供参考依据。

1　传统听评课评价模式存在的局限性

1.1　评价标准主观随意，缺乏客观证据支撑

传统的听评课最大的问题在于主观性太强。评课者只能根据自己的经验、感觉和偏好进行评课，缺乏结构性的、公认的课堂教学观察工具及评价标准，因此不同的听评者对同一堂课的评课结果可能千差万别[8]。这样的“随感”式的评课，信度较低。听评反馈意见中充满了课堂氛围不够活跃、重难点不够突出、师生互动性不够等抽象的判断，但是这些判断缺乏可供佐证的、可验证的课堂行为证据。比如，什么叫“不够活跃”？学员回答问题的次数不够多？还是学员思维参与不够多？没有数据支撑，反馈就成了个人印象的灌输，很容易引起被评教师的困惑和抵触，而不是反思和改进。

1.2　反馈内容模糊空泛，缺乏具体行为指导

正是这种主观随意性使得传统听评课的反馈环节流于形式、流于表面，听评者提出的意见是要提高驾驭课堂的能力、教学手段可再丰富一下等规律性、方向性的空对空指导[9]。这种反馈意见对一线教师而言无异于隔靴搔痒，因为不知道究竟怎么提高、怎么丰富，不是针对具体教学行为、具体教学

情境作出的反馈，没有指出做什么、怎么做，教师有心改进也无从下手。这种模糊性的反馈意见无法给教师提供清晰的可操作的“行动路线图”，不可能转化为教学行为的改进。

1.3　功能定位偏向评判，缺乏专业发展支持

更多的时候，传统的听评课被当作一种行政管理和考核手段，追求的是“分出高下”。这种“审判式”定位导致听评氛围的紧张对立，评者居于“法官”和“侦查员”的位置，教者处于“被审判”的戒备状态[10-11]。教师关心和思考的问题不再是教学问题，而是“怎么做”才能不出错，甚至为了迎接评课进行“预演”。当听评课和绩效考核、职称评定等利益分配挂钩后，其作为“形成性评价”的教师专业发展功能便不复存在。教师不敢分享自己的真实困惑，不敢提出自己的真实不足，由于双方缺乏互信，听评课就失去了教师专业发展助力的基本意义和价值。

1.4　过程设计孤立割裂，缺乏系统性改进循环

传统听评课往往是一次性的独立活动，大多存在有“听”无“评”、“有评”无“定”、“有定”无“续”等问题。上课前不会与执教者商量，课后是否有跟进措施2个环节缺位。没有课前会议，评课者心中无执教者的教学设计和意图，评课中可能出现偏离；没有课后跟进，改进措施谁来执行，执行后效果如何不得而知[12]。不能形成一个“诊断—反馈—改进—再诊断”的螺旋式上升的教师成长和评价改进过程，长期效应可想而知。

2　循证式听评课的特征

2.1　以客观替代主观

要求评课者运用结构化观察工具（如临床教学行为编码、针对某种教学行为的核查表）去采集客观、量性的课堂数据，比如，准确记录教师开展床边教学时用于讲授灌输、示教操作、启发推论、给出反馈所耗费的时间占比，或者提问中用于回忆识记的“低层次问题”与用于临床鉴别诊断的“高认知问题”所占比重，作为教学客观数据，让反馈内容不再是“我觉得……”有了不争的事实证据。

2.2　以标准替代随意评判

要求整个听评课过程必须依据教育研究与实践所确立的公认有效教学“金标准”或临床教学能力模型。医学教育中的“金标准”可以是“梅里安的七条教学原则”，或者是更医学化的“临床教学核心能力模型”（如设定学习环境、有效演示、直接观察、提供反馈等）。按照这些标准来评课，就相当于找到了诊断的理论根据，使得反馈方向正确，内容规范，避免了评课者个人偏好带来的偏颇。

2.3　以教学预后与成长方案替代终末判决

循证式听评课不是为了一堂课下个“优良中差”的最终诊断，而是基于教师后续可持续发展和教学能力提升，也就是“预后”。评课者由法官转变为“教练”或“导师”，核心任务是基于发现的教学问题，和教师共同制订一份个性化的、阶梯式的发展行动计划，是充分赋能教师、激发教师自我反思和改进的建设性评课。

2.4　以“教学多学科诊疗团队”替代“单向审视”

循证式听评课主张的是平等、互信、双向的专业对话。评课者不是高高在上的宣判者，而是与执教者共同组成一个临时的“教学多学科诊疗团队”，他们是共同探寻教学秘密的合作伙伴。他们共同在课前明确目标，共同在课后分享数据、引导提问（“您当时为什么这样处理？”“如果当时我们尝试……您觉得会怎么样？”），共同研判“病情”、共同商量“治疗方案”。这是确保反馈真诚接纳、化为行动的心理前提，是教学相长的最有效形式。

3　循证式听评课在住培师资培训中的实施路径构建

3.1　实施前的准备阶段

3.1.1　组建多学科听评课团队

循证听评课有效开展的最重要的环节，是有一支专业的、多元化的、协作的循证听评课团队。团队成员应该是专业及能力互补的。具有丰富临床经验的临床导师团队，可以帮助团队认定所教学内容的科学性、认定教师对学生的临床思维引领是否准确且恰当。同时，具有医学教师教育测量学能力的医学教育专家为团队提供理论和理论测量学的指导，以保证听评课的方法和工具的科学性。同时，非常重要的一点是寻找与被评价者背景基本相似的同伴，这可以促进评价结果的可迁移性和同事间形成开放和关怀的同伴评议文化。

因此，要对团队所有人员进行标准化的培训，其重要目的就是要让团队成员在对评价框架、评价工具把握上达成一致，加强评价者信度，确保其行为和判断有意义、符合标准。通过集体学习评价标准和典型的教学案例，分别独立评价并讨论，不断进行评价标准校准，在关键教学环节的行为识别和评级上达成共识，从而为本项工作中公正地观察课堂、科学地反馈提供可靠的人力保障。

3.1.2　开发结构化观察工具

循证式听评课的关键在于用客观证据代替主

观印象，而结构化观察工具则是获取关键证据的主要抓手，其将复杂的教学分解为可观察、可记录、可计数的微观行为，为教学分析提供可靠的数据证据。常用的工具有2大类，通用工具适合用于大多数教学情景，如课堂评估编码系统对课堂教学中的师生互动的类型和频次进行高频率、标准化的记录。专项工具则适合于医学教育的主要场景，如利用教学行为核查表来评价床边的教学，评价小讲课的课件和演讲技巧的评定量表等，针对一种教学技能进行精准的观察分析，从而提高工具的使用效能。对于国外的量表，需要进行必要的本土化修订。本土化修订包括量具条目文化含义的转换、语义表达的准确性，以及在符合本院教学实际情况的基础上，对行为指标进行增补或修改，目的是使量具的内容效度、实用性、实践指导意义在本国教学场景下更为匹配，能够灵敏、准确地反映出教学成败的关键行为[13-17]。

3.1.3 制定明确的评价框架与标准

循证式听评课的评价表是将抽象的住培师资核心教学能力逐步分解为具体的、可观察的、可测量的行为，使得评课人和被评教师对于“什么样的教学是好的教学”有一个共同、明确的定义（表1）。听评课观察的评价表由课题组通过梳理国内外文献，如临床教学能力框架、有效教学原则等，并咨询5名医学教育专家，经2轮论证，综合分析本院优秀临床教师实际经验而设计。每次听评课观察可选择2~3个核心能力领域作为观察重点并与教师课前沟通。评课人需要将客观的、具体的事实（包括学员的原话、教师的具体提问、时间点、使用材料等）在对应“观察依据”中客观、具体地表述出来，而不是将评课者的主观判断写入表格。评课人应基于表中记录的事实在课后进行具体反馈，共同确认优势并确立具体的行动计划[18-21]。

3.1.4 营造安全的心理环境

循证式听评课能否产生实际效果，首先取决于能否营造一个安全、信赖和支持的氛围。目标是消除教师的防御机制，勇于暴露自然的教学状态、个人的困惑和缺点，让听评课不再是为了追求高分的“作秀”，而是为了解决问题的“查房”。要做到这一点，首先，明确并强调发展性目的。每次活动开始之前，包括项目启动之初，都要由教学主任郑重地重申，活动的目的是帮助教师成长，无论是过程还是结果，都不会与人事考核、绩效奖励等行政管理事务挂钩，打消教师后顾之忧。其次，作出正式的保密承诺。要求所有参与者，无论是听课者还是上课者，都要共同签署保密协议，承诺不对外泄露具体的课堂情境与讨论过程、个人意见。这可以避免教师尝试失败所带来的公开负面评价，有利于他们设计更具个性化和探索性的教学尝试。最后，用行动营造信任气氛。通过预备会议让教师参与设定目标、先行采用非评判性的、描述性的语言进行反馈。这些努力都是为了给教师一个明确而一致的信号，即评课者是以“批判性朋友”和“专业同伴”身份出现，团队是一个共同探索教学技艺、互相支持的“学习共同体”。

3.2 循证式听评课在住培师资培训中的实施

3.2.1 课前会议

课前会议是循证式听评课活动的开端，为后续的课堂观摩和评议工作奠定合作、互信的基调，确保活动是基于被评教师的改进需要开展的。会议应首先由被评教师主持，清晰陈述3个方面内容：①课堂教学目标：描述本节课中学员应达到的知识、能力或态度的改变。例如“小组成员对本节课的学习后，能够自主复述急性胸痛3种鉴别诊断的思路和关键证据。”②课堂教学设计：描述为达成上述目标设计并实施的教学活动及实施步骤和原因。例如“我计划开展‘翻转课堂’，假定小组成员已对基本概念有所掌握，因此课上大部分时间将针对1例模拟病例，对鉴别诊断思路进行层层剖析，并开展分组讨论，训练其诊断决策思维。”③个人关注点：教师提出自己期待获得哪些方面的反馈意见或存在哪些困难。比如，“我最近一直在努力提升自己提问的质量，我特别希望您关注我提问‘高阶问题’（分析、评价、综合、评价类问题）的有效个数。”“我一直感觉后半节课学员的参与度不高，但一直不得其因，我特别希望您能帮我诊断分析一下原因。”

根据教师的描述，评课者与教师协商本次课堂观察的1~2个焦点问题。放弃全面“肢解”一堂课，选择一个真正有利于教师改进的突破口。比如，大家商定本次观察重点关注教师提问变化与学员认知深入发展、教学指令具体与学员完成任务等。观察焦点确定后，课中观察所使用的观察工具就会与之对应。如果观察焦点是提问变化，那么课中观察时评课者就会使用提问类型记录表；如果观察焦点是学员参与，评课者就会使用学员参与量表。这一协商过程，本身意味着评课者从“审判者”到“合作者”的立场转换，是建立安全氛围的第一步。

3.2.2 课中观察与记录

课中观察是循证式听评课收集一阶证据的关键环节。评课者在课中阶段的评课任务主要是用

表 1　住培师资核心教学能力行为指标评价表

核心能力领域	具体能力维度	可观察、可测量的行为指标(示例)	观察依据(证据来源)
教学设计与内容精熟	目标与内容	目标明确:在授课或带教开始,能够明确阐述具体的、可以衡量的学习目标; 内容正确:提供精准的信息,并且最好是与最新的指南、研究进展或文献相链接; 临床应用:理论知识与临床的完美结合,如何进行临床应用	教学计划、课件、直接观察、学员笔记
	结构与组织	逻辑清晰:内容组织严密,层层递进,重点明确,前后衔接自然; 时间把控:合理分配各个环节的时间,保障重点完成充分,按时完成	教学录像、直接观察
教学策略与互动	方法多样性	灵活:能根据教学任务灵活选用案例教学、模拟训练、小组讨论、床边教学等方法; 激发主动性:可利用设问、设置临床悬念等方法,激发学员的探究欲望和学习兴趣	教学录像、直接观察、学员反馈
	师生互动	有效提问:善于提出阶梯性问题(是什么? 为什么? 怎么办?),引导纵深思考; 回应反馈:善于对学员的回答予以评价(例如:同意、追问、澄清),鼓励学员提问	教学录像、直接观察
临床带教与思维引导	床边教学组织	流程规范:能严格遵守“床旁问诊+示教—学员操作—下床小结”规范流程,尊重患者; 实时指导:对学员的操作进行实时、具体的指引和提示	直接观察、学员访谈、患者反馈
	临床思维引导	思维的可视化:要求学员口头或文字描述诊断的推理过程、鉴别诊断及理由; 错误分析:把学员的错误进行改编,变成教学样本,让学员分析错误原因,找出正确答案	病例讨论记录、教学录像、学员反思日志
评估与反馈	形成性评估	方法多样:灵活运用 Mini-CEX、DOPS 对学员进行形成性考核; 以评促学:评估结果可以直接作为下一步的个性化教学指导的切入点	评估表格、教学记录、与学员的沟通记录
	有效反馈	具体及时:能针对具体的行为表现,能及时做出回应; 建设性:能根据标准找出不足,同时给出 1~2 条针对性、可操作性的整改意见	直接观察、反馈会话记录
学习氛围与职业素养	学习氛围营造	安全环境:学员名字记得住,允许提问,包容学员的犯错,把犯错当成学习的机会; 公平正义:让学员们都有公平竞争和展示的机会	直接观察、学员访谈、问卷调查
	职业角色榜样	以身作则:语言中体现同情病人、工作负责、坦率、认真等; 团队精神:与其他学科专业人员在交叉学科背景下合作,对学员进行团队精神的讲解	直接观察、患者反馈、同行评价
沟通能力	师生沟通	表达清楚:语言表达清楚、准确、有逻辑、容易理解; 积极倾听:学员发言时,要集中精力,通过点头、重复等表现形式表示在听	直接观察、教学录像
	医患沟通示范	尊重病人:当着学生的面讲如何取得知情同意,如何保护病人隐私,如何用通俗的语言把病情解释给病人; 人文关怀:即感同身受,对病人及家属的情感需求表示理解,给予情感支持	直接观察、标准化病人反馈
因材施教	个性化指导	能力评估:能够通过观察或询问,迅速掌握不同学员的不同基础和偏好; 差异化教学:可为不同能力的学员提供不同的目标、不同的挑战	教学计划、与学员的互动记录、学员成长档案

课前会议选取的工具,客观、中立地记录课堂上真实发生的事实。评课者必须保持中立,记录的是“是”,而不是“应该”。也就是说,记录时尽量不使用带有评价性词语的描述,如“好/不好”“有效/无效”“应该/不应该”等。例如,有的学员没有参与讨论,评课者不能记录为“该学员没有认真听讲”,而是客观描述“在 20 min 的小组讨论中,学员 A 一直是低头不说话,没有与组员的眼神或语言交流”。这样的描述保证了证据的纯粹性,为课后基于证据的理性分析提供了条件。

记录真实现场的证据是定量与定性的结合,定量的目的是寻找教学行为结构,评课者用量化工具(结构化观察表/清单)记录某种教学行为的次数,这种记录行为的时间占用量等。如提问技术:聚焦“提问的认知”水平证据记录表是 1 个两列表格,每当教师有 1 次提问,就在列相应的“提问的水平”(低水平:记忆/理解 vs. 高水平:应用/分析/评价/创造)画一个“正”字。课后再记录两者比例。如“这节课教师 15 次提问,10 次低水平(67.00%)提问,5 次高水平(33.00%)提问”。再如时间结构:教师用时间日志记录下本节课自己理论讲解、案例讨论、示范操作各用了多少时间,客观呈现结构重心。

定性记录记下教学情境的细节和内涵。评课者尽量记录教学事件中关键环节的对话、手势、表情、情境等。如反馈方式:针对焦点“反馈是否具体”,不是只记录“教师进行了反馈”,而是应该记录

下他说了什么，做了什么。如："10:15，对学员 B 的穿刺操作进行反馈：'你刚才的进针角度很好，但是忽视了铺巾的范围，这样就会把这个角污染了（手指位置）。下次铺巾的时候，可以先把巾铺开，保证铺巾范围在操作区域内。'并做手势比画。"

3.2.3 课后反馈与研讨

课后反馈与研讨是课中获取的客观性证据，通过一次结构化专业对话，转化为被评教师的自我觉察与具体行为。在会议开始，评课者以不评价、不预测的方式，借助图表、时间线、原始笔记摘录或者简单的数字统计等可视化方式呈现课中记录的关键观察数据。为后续的对话提供一个客观、共享的事实起点，避免对话中分歧来自感受与印象。如"王老师，关于'提问类型'的观察，是我们课前约定的关注点，数据如下：在 30 min 的讨论中，你总共提了 12 个问题。其中有 9 个事实回忆型封闭式问题，例如'这个病的诊断标准是什么'，3 个要求分析推理的开放式问题，例如'你为什么首先排除这个诊断'。"这是客观的记录。

数据呈现之后，评课者要从"呈现者"角色切换到"引导者"，通过一系列引导式提问，启发教师自己分析和解读数据。可以从以下几个角度提问：①感知与发现："你看到这些数据的时候，你看到了什么？""你有什么出乎意料的发现？"②意图与决断："你当时为什么选择以这样的方式（高密度封闭式提问）来推进讨论？"③影响与后果："你认为这样的提问方式对学员思维深度、参与度有什么影响？"④替代与可能性："假设我们当时将其中一个事实性问题替换成应用性问题'你准备如何为这位患者做一个长期的管理方案'，课堂讨论会有什么变化？"

在教师反思的基础上，评课者适时引入教育学理论和有效教学原则，将对话从"发生了什么"提升到"为何发生"以及"如何优化"的层面。可以结合数据和理论，共同分析教学行为的有效性，并探讨替代策略。如"您刚才提到希望提升学员的临床推理能力，这与我们倡导的支架式批判性思维是一致的。数据显示我们目前多以封闭性提问搭建基础，下一步或许可以尝试在几个关键节点，像我们刚才设想的那样，嵌入 1~2 个高阶问题，为学员的思维'搭设脚手架'。我这里有几个其他科室常用的提问策略，我们可以一起看看是否适用。"这样共同制订一个具体、可行的短期行动计划改进计划，这是将反思转化为行为改变的关键。如"在下周的 2 次小讲课中，每次至少设计并提出 3 个要求学员应用、分析或评价的开放性问题。我会将这个问题设计环节明确写入我的备课教案中。您可以分享刚才提到的高阶问题库作为参考。2 周后，我们花 10 min 简短交流一下实施后的感受和学员反应，您看如何？"

4 实施成效分析

4.1 研究对象

本研究对象为北京清华长庚医院国家级住培基地内自愿参加循证式听评课项目的 36 名住培师资。采用目的抽样方法，选取涵盖内、外、妇、急诊、耳鼻喉科等主要科室，职称涵盖高、中、初级，以及不同教龄（从新任导师到资深导师）的住培师资，且自愿参加，知情同意。研究对象全部参与了项目，接受所有干预和数据收集，无退出和失访。纳入标准：①研究期间属于清华大学附属北京清华长庚医院国家级住培基地的住培带教师资，并承担了实际带教任务者；②自愿签署知情同意书，并愿意全程参与为期 1 年的循证式听评课项目以及参与所有培训、观察和反馈活动者。排除标准：①研究过程中有长期外出（如参加进修、援教、产假等）计划，无法保证完整投入整个项目周期的研究对象；②同时参与其他可能影响教学行为且有系统性影响的教学改革项目，无法保证教学行为相对独立、一致的研究对象。

研究前期（1 月至 2 月）：对所有研究对象实施问卷前测和半结构化访谈，了解研究对象教学信念、自我效能感、传统听评课认知等，收集基线数据。中期（3 月至 10 月）：实施 2 轮完整循证式听评课（每轮由准备阶段、课前会、课中观察、课后反馈 4 个部分组成）。同时，通过参与式观察记录下反馈会关键互动环节，并收集所有观察工具原始数据。每轮结束后，对部分教师进行小组访谈，了解其即时体验与困惑。后期（11 月至 12 月）：对所有研究对象实施问卷后测和深度访谈，并收集同期学员评教分数。对比数据，分析变化。对所有质性资料进行编码、主题分析。

4.2 评价方法

4.2.1 定量评价工具

采用《住培师资教学自我效能感量表》评估教学内在信念与主观信心。在 Tschanen-Moran 等编制的《教师自我效能感量表》基础上，结合住院医师规范化培训的教学内容进行本土化修订。修订后的量表包括教学设计与准备、教学策略与实施、评估与反馈、学习氛围与沟通、循证教学与专业发展 5 个维度，17 个条目。采用李克特 10 点计分，从"完全没信心"至"完全有信心"分别计 1~10 分。分值越高，

提示其该项教学能力自我效能感越强。总分范围为17~170分，各维度得分可单独计算，分别分析不同能力领域的信心变化。量表总体内部一致性信度系数Cronbachs's α值为0.92，各维度α值为0.78~0.88，提示具有较好的内部一致性信度。内容效度由5名医学教育专家评定，各条目I-CVI均>0.78。

采用《教学实践行为频率量表》评估教学实践外显、可观察的教学行为发生频率。《教学实践行为频率量表》由研究者根据临床教学核心能力框架自编，用于评估教师日常教学实践中关键行为的发生频率。包括教学目标与教学设计（3个条目）、教学策略与临床思维引导（4个条目）、形成性评估与反馈（4个条目）、学习氛围营造与职业素养（4个条目）、教学反思与循证实践（3个条目）5个维度，共18个条目，采用李克特5点计分，从"从未"到"总是"依次计1~5分。分值越高，表示该教学行为在教师日常实践中的发生频率越高。总分18~90分，可用于评估教师教学行为的总体改善情况。《教学实践行为频率量表》总的Cronbachs's α系数为0.89，重测信度为0.85，量表整体具有良好的内部信度。内容效度由专家评议和预测试保证，使条目能够有效反映住培教学的关键行为。

采用《住培学员对师资教学效果评价表》从学员角度评价教师教学效果的改变，该表系本院教学管理部门长期使用的评价量表，量表包括教学目的和组织条理（3个条目）、临床思维指导程度（3个条目）、点评和考核的有效性（3个条目）、学习氛围和交流（3个条目）、教学投入和整体教学效果（3个条目）5个模块，共计15个条目。量表采用李克特5分制评分，从"非常不同意" 到"非常同意"依次计1~5分。分数越高，代表学员对教师的教学效果评价越高。分数范围15~75分。该量表内部一致性信度系数Cronbachs's α稳定在0.88以上，稳定性良好。量表条目从临床教学关键目标出发，内容效度良好，且经过多轮教学督导验证。

于项目中期6个月（T0）、项目结束时（T1）分别对36名师资发放《住培师资教学自我效能感量表》及《教学实践行为频率量表》进行问卷调查，同时收集2023年度（干预前）、2024年度（干预后）学员填写的《住培学员对师资教学效果评价表》，分别收集有效问卷79份、88份。数据采用SPSS 26.0进行分析，对师资前后测数据采用配对样本t检验，对不同年份学员评价采用独立样本t检验，检验水准α=0.05。

4.2.2　定性评价方法

半结构化访谈分别在项目实施前后，对18位参与教师采用一对一的深度访谈，访谈时间30~45 min/人，主要了解教师的教学反思程度、接受反馈的程度、对自身的专业认同、对循证听评课的体验和认识等，访谈录音转写成文字资料。

焦点小组访谈：在项目中期进行2场焦点小组访谈，8~10人/场，时间60 min左右，了解教师共同的感受、共同的问题和对学习共同体的感知。由研究人员担任主持人，设计提纲，进行半结构访谈。

系统收集3种文本：①教师个人反思日志（教师每参与一轮活动后提交1篇，共计72篇）；②课后反馈会议记录（共计36份）；③教师共同拟定的教学改进计划（共计36份），通过分析文本数据来考查教师的反思内容与反思方式以及教学行为改进的真实动态过程。所有质性资料使用主题分析法进行资料分析与整理。第一步：由2名研究者分别阅读所有文本资料，然后进行初步的资料编码。第二步：研究者协商，合并编码，形成初步的主题。第三步：反复比较、提炼和修改，形成最终的主题。第四步：对于不同的地方，讨论至达成一致。通过质性分析来寻求对研究问题的一种解读，把握教师的心理变化、反思方式变化与行为改进的脉络。

4.3　初步成效

4.3.1　住培师资教学自我效能

为了解循证式听评课对师资教学自信的影响，分别于项目中期（T0，6个月）及项目结束时（项目结束时，T1）使用《住培师资教学自我效能感量表》进行测评。结果显示（表2），听评课师资教学自我效能感总分显著提升，从维度平均分来看，提升的最高分为循证教学与专业发展、评估与反馈2个维度下的条目，即对于可训练、可观察的具体教学技能，如使用结构化的观察检查表或者其他观察办法客观地分析自己的教学、基于明确的标准给出描述性的具体反馈、有意识地提出不同认知层次的问题等条目；对于需要更多经验积累和情景判断的综合性能力等条目效能感得分提升处于中等。但是，面对"需为不同背景的学员准备不同的教学材料及案例""能否有效调动没有学习兴趣或缺乏自信心的学员的积极性"这2项更高的要求，师资的自我效能虽然也有小幅提升，但差异没有统计学意义。

4.3.2　教学实践行为频率量表

对教学实践行为发生频率的测量发现（表3），接受师资行为发生频率明显提高。其中增长幅度最大的为教学策略以及临床思维引导和形成性考核与反馈维度。

表 2　住培师资教学自我效能感比较（n=36）

维度与具体条目		T0（开展 6 个月）	T1（项目结束）	t 值	P 值
教学设计与准备	我能够针对每次临床带教过程制定明确的、可以量化的学习目标；	7.25 ± 1.52	8.85 ± 0.93	7.12	<0.001
	我可以根据学习的课程内容，设置有条理、有重点的教学步骤；	6.92 ± 1.61	8.28 ± 1.05	5.89	<0.001
	我能够为不同基础的学生，准备不同的教学资料和案例	5.45 ± 1.78	6.02 ± 1.65	1.98	0.055
教学策略与实施	我有意识地提出不同认知水平的问诊问题，以启发学员的临床思维；	5.88 ± 1.70	7.95 ± 1.28	8.45	<0.001
	我能够指导学员总结和表达鉴别诊断和做出诊断的依据；	6.35 ± 1.58	8.15 ± 1.22	7.33	<0.001
	当学员有不理解的地方，我能用多种方法进行讲解；	7.05 ± 1.48	7.52 ± 1.35	2.05	0.048
	我能合理安排教学时间，做到该讲的讲足	6.58 ± 1.65	7.25 ± 1.41	2.92	0.006
评估与反馈	我能根据具体明确的标准，给出具体的、描述性的评价；	5.72 ± 1.82	8.22 ± 1.18	10.15	<0.001
	我可以提供批评性反馈而不让他们感到被打击；	6.18 ± 1.75	7.08 ± 1.52	3.45	0.001
	我能合理有效地运用 Mini-CEX、DOPS 等工作形成性评价工具；	6.85 ± 1.70	8.05 ± 1.30	5.01	<0.001
	我可以将评估结果转化成下一步具体的教学行为	5.95 ± 1.66	6.88 ± 1.44	3.78	<0.001
学习氛围与沟通	我能创设安全、尊重、敢问的课堂情境；	7.42 ± 1.55	8.35 ± 1.12	4.02	<0.001
	我能有效地激发学习兴趣不高、信心不强的学员；	5.82 ± 1.79	6.15 ± 1.62	1.25	0.220
	我能够很清楚地给学员讲授复杂的临床问题	7.28 ± 1.46	7.75 ± 1.25	2.22	0.033
循证教学与专业发展	我可以使用结构化的观察工具 / 方式对执教过程进行客观的审视；	4.55 ± 1.85	7.25 ± 1.48	11.02	<0.001
	我会根据教学反馈结果，制定出有针对性、可操作性的个人发展计划；	5.75 ± 1.72	7.68 ± 1.33	7.85	<0.001
	我能够运用医学教育中新的知识来指导我的教学实践	5.38 ± 1.81	6.05 ± 1.59	2.58	0.014
效能感总分		6.21 ± 1.18	7.54 ± 0.89	9.87	<0.001

表 3　教学实践行为频率量表前后测数据对比分析（n=36）

维度	条目	T0（开展 6 个月）	T1（项目结束）	t 值	P 值
教学目标与教学设计	在进行临床带教前就要向学员明确学习的目标；	3.25 ± 1.05	4.32 ± 0.65	7.89	<0.001
	教学流程经过精心设计，环环相扣；	2.95 ± 0.98	3.88 ± 0.74	6.54	<0.001
	根据教学目标，提前准备特定的教学材料	3.45 ± 1.02	3.72 ± 0.91	1.85	0.073
教学策略与临床思维引导	有意识地提出开放性的、能够促进高阶思维的问题；	2.12 ± 0.92	3.78 ± 0.81	11.45	<0.001
	引导学员系统地陈述鉴别诊断及依据；	1.85 ± 0.83	3.25 ± 0.92	9.80	<0.001
	运用“思维可视化”法帮助学员梳理临床信息；	1.68 ± 0.78	2.05 ± 0.85	2.55	0.015
	根据学员反应，灵活调整教学进度与深度	2.68 ± 0.87	2.82 ± 0.91	1.12	0.271
形成性评估与反馈	观察学员操作后，立即提供反馈；	2.35 ± 1.02	4.02 ± 0.73	12.03	<0.001
	给予具体、描述性的反馈，而非模糊评价；	2.28 ± 0.95	3.85 ± 0.76	10.89	<0.001
	与学员共同商讨具体的后续改进策略；	2.48 ± 0.98	2.95 ± 1.02	2.89	0.007
	有目的地使用 Mini-CEX、DOPS 等测评工具	2.88 ± 1.12	3.92 ± 0.82	6.35	<0.001
学习氛围营造与职业素养	鼓励并等待学员提问，并耐心解答；	4.15 ± 0.78	4.52 ± 0.55	3.65	<0.001
	营造让学员敢于暴露不足的安全学习环境；	3.75 ± 0.89	4.28 ± 0.71	4.82	<0.001
	刻意示范表达同情心、尊重、获得知情同意	4.05 ± 0.82	4.35 ± 0.68	2.65	0.012
教学反思与循证实践	课后简单回顾教学，思考优劣与改进点；	3.82 ± 0.95	4.25 ± 0.74	3.42	0.002
	参考评价结果或观察数据规划教学改进；	2.15 ± 0.88	3.08 ± 0.92	6.18	<0.001
	尝试将新的教学理念或方法应用于实践	2.55 ± 0.90	2.88 ± 0.87	2.25	0.031
行为频率总分		2.76 ± 0.65	3.58 ± 0.58	9.12	<0.001

4.3.3　住培学员对师资教学效果评价

结果显示（表 4），2024 年住培学员对师资教学效果评分的总体分数显著提升。教学核心技能更结构化，在教学的目标与组织是否清晰和教师引导临床推理是否深入条目，学员认为教师开课之初对教学目标介绍更清晰，讲授或演示内容逻辑层次更突出，可提供更多蕴含开放思维、值得反复推敲的问题，临床推理过程导向性更强。反馈更及时具体、学习氛围和教师投入感更强烈。数据还表明，在与学员共同商讨改进计划、有效掌控教学进度、有效引导学员纠正思维偏差、关注学员个性化需求上，虽然 2024 年的评价平均值有一些提升，但并不具有统计学上的意义。

4.3.4　访谈与反思日志分析

通过将半结构化访谈的转录文本和教师反思日志进行仔细的主题分析，整理出了 3 个主题，充分

表 4　住培学员对师资教学效果评价的对比分析

评价维度	具体评价条目	2023 年（n=79）	2024 年（n=88）	t 值	P 值
教学目标与组织清晰度	在带教前，教师会告诉我们这一次要学什么；	3.82 ± 0.85	4.48 ± 0.62	5.12	<0.001
	教师讲得清楚，提纲挈领，我跟着他走，一点也费不着劲儿；	3.95 ± 0.78	4.42 ± 0.59	4.35	<0.001
	教师能够合理控制教学进度和教学时间，确保重点内容能够得到充分讲解	3.75 ± 0.81	3.92 ± 0.74	1.45	0.152
临床思维引导深度	教师给出的题目能够让我思考，而不是单纯地去背诵；	3.52 ± 0.90	4.25 ± 0.68	5.78	<0.001
	在分析病历时，教师会指导我理清鉴别诊断以及各自的依据是什么；	3.65 ± 0.88	4.30 ± 0.66	5.24	<0.001
	当我自己在思考上出现失误时，教师能够及时指出错误，并引导我往正确的方向上思考	3.58 ± 0.92	3.75 ± 0.85	1.12	0.265
反馈与评估的有效性	教师能在我进行操作后，及时给予反馈；	3.42 ± 0.98	4.35 ± 0.65	7.15	<0.001
	教师给的反馈也很具体，总是能告诉做得对/不对；	3.35 ± 0.95	4.28 ± 0.64	6.88	<0.001
	教师在对我的不足进行批评的同时，往往会提出中肯、可行的建议	3.15 ± 1.02	3.28 ± 0.95	0.78	0.438
学习氛围与沟通	教师鼓励我们多提问，对提出的问题会耐心解答；	4.22 ± 0.72	4.58 ± 0.51	3.85	<0.001
	在这样一位教师面前，我觉得安全，能够显露出自己操作、知识方面的不足；	3.82 ± 0.85	4.35 ± 0.67	4.52	<0.001
	教师能关注到不同学员的个性化学习需求	3.38 ± 0.92	3.48 ± 0.88	0.68	0.498
教学投入与整体效能	我能感受到这位教师对教学工作充满了热情；	4.15 ± 0.80	4.45 ± 0.62	2.89	0.005
	通过这位教师的指导，我认为自己在临床知识和技能上有显著提高；	3.88 ± 0.82	4.32 ± 0.66	3.95	<0.001
	总体而言，我觉得他是一位杰出的临床教师	4.10 ± 0.88	4.22 ± 0.79	0.95	0.345
评价总分		3.72 ± 0.70	4.18 ± 0.55	6.50	<0.001

展现出参与教师在观念和态度方面的巨大转变。

主题 1：从“恐惧防卫”到“开放接纳”，教师心理定位的转变。在访谈中，几乎所有教师都生动地描述了他们参加听评课前后迥然不同的心理状态。教师在传统听评课中认为听课是一种“审判”、一种“工作检查”，因此产生焦虑紧张，以及为了规避而“表演式教学”的心态。一个主治医师教师的发言非常具有代表性：“以前一听到要听我的课，心理就打鼓，提前一周开始‘准备’，就是想把自己最好的一面、不犯错的一面呈现出来，感觉很吃力，并且对实际教学帮助并不大。”但是在经历了循证式听评课后，教师的转变非常彻底。从“检查”到“诊断”，从“审判”到“帮助”。另一个教师说道：“现在的心态完全不一样了，甚至是‘期待’别人来听我的课，因为我知道观察者是我的‘战友’，他们是来帮我找出不足。每次课后都觉得做了一次‘教学体检’，很有收获。”这是从“恐惧防卫”到“开放接纳”最直观的结果，是所有深层次教学改进的心理基础，这是循证式听评课最重要的心理安全文化建设成果。

主题 2：教学反思从“模糊感觉”走向“证据思维”。反思日志内容鲜明体现了教师反思从模糊走向具体。项目初期教师的描述多是模糊的、感觉式的，比如“今天课堂气氛有点沉闷”“学员似懂非懂”。他们普遍感觉“有问题”，但说不出，比如一位教师的日志写道：“总感觉自己的课有明显可以改进的方面，但就是不知道到底哪里出了问题，感觉像隔了层毛玻璃，看不清楚。”随着项目的进展，教师反思日志中开始出现大量具体的、数据式的描述，如一位教师在日志中反思：“通过观察记录，我发现整节课问了 20 个问题，但 18 个都是是不是、对不对。我一直在赶节奏，完全没有留给学员思考的空间。数据让我看到了巨大差距。”

主题 3：教学行为改进从被动接受到主动设计。这一主题反映了教师在教学设计以及实施中能动性的提高。教师们的访谈中，已经不再将自身定位为教学大纲内容的被动承接者，而成为了教学的“主动设计者”，他们的谈话中开始出现“我尝试了……”“我下一步尝试……”等主动语态。一位教师的表述可以充分说明这个问题：“我现在的备课状态下，会有一种下意识：我这里是不是可以设置成一个开放式病案情景，我今天准备提问的问题，是一个可以启发他们临床思维的问题，还是一个单纯的提取知识点的提问？这种‘设计感’是以前所没有的。”反思日志记录了这种主动性转变为实际行动的过程，不少教师的日志中不再仅有问题，更有针对性的、循证的“行动处方”。如：“根据上次反馈，本周我在 3 次床边教学中尝试刻意练习‘总结性反馈’技巧。我逼着自己非得引用学员的一句话和一件事不可，确实效果更好了，学员的认可度提升了。”从接受旁观者“他者”的建议，进入到基于“自悟”和“自断”的“自驱”阶段，教师真正实现了教学改进的自我主导、自我赋能。

4.3.5 焦点小组访谈与文档分析

4.3.5.1 专业学习文化的形成

焦点小组实录揭示了一个强大的“专业共同体”正在凝聚。研究者发现，教师们谈论课改，说话方式发生了明显变化，他们越来越多地使用“我们”“咱们科室”“一起研究”等表示集体意义的词语，过去常用的高调个人责任意识的话语“我”或者指向管理部门的话语“他们”，则很少出现了。这种话语体系的变化说明，教师开始从心理将自己界定为一个为共同发展努力的专业共同体。此外，更为重要的是同行之间专业对话内容和水平的提高。他们不再是“课堂气氛好不好”，或者“学员表现怎么样”这类印象式、表层式谈论，教师谈论的是教学技术和理论在实际教学中的具体运用。小组成员会讨论“如何把写病历的SOAP结构（主观、客观、评估、计划）运用到对学员操作情况反馈中”、如何运用“一分钟预科教师”模型中“规则化教学点”这一步骤时采用的不同情况下的策略运用。这种共同体真正发挥作用，需要建立在共同认知图式基础上的专业化的深度对话，使教和学的智慧流动于组织，沉淀于组织。

4.3.5.2 教学改进的闭环管理

对反馈会议记录和教师个人改进计划的档案研究揭示了作为持续专业学习的循证式听评课运作机制。这一档案文件形成了完整的可供追踪的“病史”，从发现问题到行为改进形成了一个完整的闭环。在1份档案中清晰地描绘出其进化路线图，反馈会议记录是“联合诊断”，里面记录着诸如此类的对话。听评者：“数据表明您在20 min的讲课中，没有停下来检查学员的掌握情况。”被评者：“原来我有这方面的顾虑，担心停下来，打断思路，但看来学员也听不懂了。”而跟进的个人改进计划是“联合处方”，被评者制定出高度具体的行为改进目标，例如：“下周内科小讲课，我将争取在讲完‘心衰的病因’和‘治疗原则’这两大段落后，主动检查学员的掌握情况，方法是让学员快速复述或问要义。”尤其重要的是，后期的随堂听课证据正是该教师“行为跟进与追踪”的印证，证明其的确在接下来的教学中使用了该方案。而这样的“证据诊断→协商处方→行为跟进→成效检验”闭环，不仅意味着循证型听评课在教师个性化成长方面的“对症下药”，更意味着这套“计划部署—组织实施—成果督导—再部署”的持续改进文化被组织制度化地融入教师日常的教学生活之中。

4.4 教师行为改变的关键因素

对访谈、反思日志进行质性分析，梳理出教师行为变革的主要发生过程有以下3个环节：①建立心理安全：通过签署保密协议、明确发展性定位等，减少教师的阻抗行为，由“演戏”回归“教学”。②基于证据对话：借助结构化工具，反馈有事实依据、具体可信，教师认可并行动。③建立反馈闭环：“课前—课中—课后—回访”四位一体，改进计划有效落实并动态调整。对结构化反馈的回应程度与教师职称、教龄有关：教龄长、职称高的教师更关注教学策略的更新，教龄短、职称低的教师更关注基本教学技能的强化。

5 讨论

5.1 循证式听评课实施的关键要素分析

本次研究使得循证式听评课评价体系有3个关键内涵：一是以“证据”替代“经验”，这是质变。盲目与随意是听课评课原本的问题。“循证”研究设计结构化工具、厘清评价标准，使得抽象的教学能力与素养指标被分解为一个个可以看得见的行为指标；使课堂对话建立在诸如“高阶问题占比”“学生反馈具象化”等数据基础之上，为改进指明一个个抓手。二是以“发展”替代“评判”，这是保证。正是由于本研究在听评课评价结果与教师的业绩考核脱钩、提前签署保密承诺书、对话建立在心理安全之上等种种努力，才最终真正将教师的上课心态由“防范—作秀”向“开放—求索”转变。只有当教师相信活动本身就是为了帮助自己变得更好，才会把自己的问题呈现出来、把问题当成问题、做真正的挑战、做出真正的改进。最后，“系统性的闭环”取代了“孤立性的事件”。本文所构建的“会课前协商—会中跟进—课后反思—持续改进”提供了教师改进的“闭环路径”。尤其是教师聚在一起会前协商好、确定好观察要点，会后又一起商量出“改进处方”及计划，听评课由此变为一个“持续性”的帮助式教学改进的专业活动，正是这一机制在发挥作用，让真正的改变成为可能[22-24]。

5.2 循证式听评课的先进性与创新性

循证式听评课不是对传统听评课的修修补补，而是一种思想变革、工具变革和目的变革。思想上，完成了从“经验”到“循证”的转变。循证式听评课将“循证”这一医学教育领域的核心思想移植到听评课领域，使师资培养根植于客观证据、科学共识。循证式听评课将规范化的观察工具、标准化的

评价维度引入听评课，将听评课从“感觉的艺术”转变为“观察的科学”。循证式听评课将评课者从“法官”转变为“教练”，将听评课课堂从“审视”转变为“教学 MDT”，真正激活了教师作为主体的生长力。教师反思从“模糊感觉”到“证据思维”的提升，行为改进从“被动实施”到“主动设计”的跨越，这是传统模式所不能比拟的。

5.3 研究的局限性

虽然本研究取得了一些成绩，但仍存在如下不足：①研究对象来自单一的国家级培训基地，又为自愿参加研究，样本较少，并存在着自愿参加研究时的偏倚，其结论的外部效度有待商榷。需要在不同地区、不同级别的培训基地进行大样本检验。②观察时间为 1 年，尽管观察到一定的即时效应，但循证式听评课对教师教学行为的稳定性，后续住培学员走向工作岗位的胜任力及对住培基地教学文化的深层影响等还需长期追踪观察，有待验证。③该模式的真正有效实施需要多专业合作，对听评课的观察者的培训及一定时间的投入，这对于不少培训基地而言是难以实现的。不过，如何在不违反这些基本原则的前提下，发掘出成本低廉的、可贯彻的推行方式，确实是需要事先考虑的实际问题。至于本文中发现效果不显著的项目（例如启发分化式教学、对“厌学”学员施加动力），则更应在日后的师资培训里开辟专业学习，并尽量为教师寻找长期的、经常性的、专业的引领者及同伴分享者。

利益冲突　所有作者声明不存在利益冲突

作者贡献声明　秦贺、李欣：确定研究题目、设计研究方案、撰写及修改论文；陈雯婧：资料收集；叶京英：提供研究建议、审校论文

参考文献

[1] 罗璐瑶, 刘岚, 赵秋, 等. 住院医师规范化培训指导师资准入评价指标体系的构建[J]. 医学教育管理, 2024, 10(6): 735-740. DOI: 10.3969/j.issn.2096-045X.2024.06.018.

[2] 张黎峰, 张婕, 毛雯靖, 等. 国内外住院医师规范化培训师资胜任力研究进展[J]. 中国社会医学杂志, 2023, 40(3): 277-280. DOI: 10.3969/j.issn.1673-5625.2023.03.008.

[3] 廖伟, 刘淼, 毛玉琳. 循证式听评课：为何循证、循何证、如何循证[J]. 中国教育学刊, 2023(4): 86-91.

[4] 吴南中, 李少兰, 陈明建. 人工智能支持的教师循证教育：理论架构与行动网络[J]. 电化教育研究, 2023, 44(5): 36-43. DOI: 10.13811/j.cnki.eer.2023.05.005.

[5] 陈丹, 詹静, 孙绵涛. 修身成己之于教育家型教师内生性成长的意蕴研究[J]. 教师教育研究, 2025, 37(6): 13-19. DOI: 10.13445/j.cnki.t.e.r.2025.06.016.

[6] 宋国语, 黄嘉莉. 中国教师培养体系的结构性矛盾、成因与优化：基于复杂系统论的系统综述[J]. 教师教育研究, 2025, 37(3): 47-55. DOI: 10.13445/j.cnki.t.e.r.2025.03.007.

[7] 陈韶华, 郑敏. 论临床教学中的提问与反馈技能[J]. 医学与哲学, 2025, 46(15): 72-76, 80. DOI: 10.12014/j.issn.1002-0772.2025.15.15.

[8] 李倩, 高屹琼, 刘薇. 高校青年教师课堂教学能力提升研究[J]. 教育教学论坛, 2023(15): 33-36.

[9] 覃千钟, 陈灵华. 教师教学技能实践的反思与重构：基于反思性实践论的视角[J]. 教育理论与实践, 2024, 44(17): 25-30.

[10] 李建锋. 从评课视角观测教师学科教学能力[J]. 物理教学, 2023, 45(10): 29-31, 28.

[11] 毕建文. 教师听评课存在的问题及改进策略[J]. 青海教育, 2024(12): 48. DOI: 10.3969/j.issn.0529-3502.2024.12.032.

[12] 欧阳鹏. 高校教学督导工作态势、问题及对策：基于“势道法术器”的视角[J]. 大学教育科学, 2024(3): 63-72. DOI: 10.3969/j.issn.1672-0717.2024.03.08.

[13] 汤少冰, 杨兰, 陆佳芳. 人工智能辅助听评课的应用研究：基于与传统听评课方式的比较[J]. 广东第二师范学院学报, 2024, 44(6): 70-82.

[14] 陈昭稳, 左荣欣. 基于云端高校教学督导听课评价系统数据库建设研究[J]. 经济师, 2023(12): 187-188. DOI: 10.3969/j.issn.1004-4914.2023.12.090.

[15] 汪维富, 毛美娟, 闫寒冰. 精准教研视域下的教师评课反馈分析模型研究[J]. 电化教育研究, 2022, 43(1): 122-128. DOI: 10.13811/j.cnki.eer.2022.01.016.

[16] 邰枫, 袁伊航, 李振泉. 运用正态分布理论的教学督导课堂评价研究[J]. 创新创业理论研究与实践, 2024, 7(9): 21-24.

[17] 郭亚东, 秦兰英. 以评课指标体系为目标导向的高校课堂教学探讨[J]. 大学教育, 2023(10): 55-57. DOI: 10.3969/j.issn.2095-3437.2023.10.016.

[18] 汪茂华, 马跃, 张慧琳. 教师创造性教学行为对学生创造性问题解决能力的影响：内部动机的中介作用[J]. 教育学术月刊, 2024(7): 81-90. DOI: 10.16477/j.cnki.issn1674-2311.2024.07.006.

[19] 卜凡丽, 唐烨伟, 赵一婷, 等. 跨学科教学能力测评事理图谱：内涵、表征与模式[J]. 电化教育研究, 2024, 45(11): 108-114. DOI: 10.13811/j.cnki.eer.2024.11.015.

[20] 陈雅川, 高宏钰. 情境取向的教师教学能力：内涵要素与发展路径[J]. 教师教育研究, 2024, 36(1): 42-48. DOI: 10.13445/j.cnki.t.e.r.2024.01.013.

[21] 方海光, 洪心, 舒丽丽, 等. 基于课堂智能分析大模型的教师教学能力分析框架及其应用研究[J]. 现代教育技术, 2024, 34(2): 43-52. DOI: 10.3969/j.issn.1009-8097.2024.02.005.

[22] 任雪明. “研”在“听”前：一种课堂观察新范式[J]. 化学教学, 2024(10): 20-24.

[23] 杨伊, 陈昌来, 陈兴冶. 基于多模态语料库的教师评价类话语特征研究[J]. 教师教育研究, 2024, 36(5): 22-29. DOI: 10.13445/j.cnki.t.e.r.2024.05.001.

[24] 秦呈玲. 立足“三研”助推教研组跨越式发展[J]. 中国教育学刊, 2023(7): 106.

（收稿日期：2025-12-01）

（本文编辑：唐宗顺）