【教育研究】

面向“未来战略中医临床家”的多元能力评价指标体系设计*

杨正[1] 刘希[2] 方旖旎[1] 梁雪[1] 林洁[1] 张妍[1] 李英帅[1] 王琦[1#]
(1 北京中医药大学王琦书院、国家中医体质与治未病研究院 北京 100029;
2 北京中医药大学教务处 北京 100029)

摘要:目的 将“王琦班”四维能力框架转化为多元能力评价指标体系,为“重要而难测”的中医高阶能力提供可操作的评价设计方案。**方法** 采用横断面问卷对30名专家进行探索性需求调查,揭示重要性认知与权重配置的落差;再以文本分析与专家论证为基础,依托四维框架构建体系。**结果** 专家虽将战略与创新能力列为首要改革方向,却在百分制权重配置中倚重学业绩点(GPA均值46.5分,战略/创新项目11.6分);在突出战略能力的专家中,81.8%仍赋予GPA更高权重。构建的体系涵盖学术传承、思维训练、能力塑造、价值引领4个维度,细化为16个二级指标并配置6类核心评价工具,按“能力界定—行为锚点—评价工具—计分规则—决策规则”转化链,为难以测量的高阶能力逐项设计评价内容;另有3项针对性设计:将价值底线设为门槛性指标、调整晋级权重,以及拟采用“师生独立临诊比对”,以师生辨证的一致程度(辅以治法一致性)间接测评师承中难以言传的隐性辨证能力(如在具体病例中对辨体、辨病与辨证的综合权衡)。对照“战略中医临床家”5项特征,经典根基与临床辨证覆盖较充分,理论原创潜力、战略意识与领军潜质的工具化程度依次递减,构成当前主要缺口。**结论** 该体系为“隐性辨证能力难测、价值底线易被分数抵消”等难题提供了一条待验证的评价设计思路;其内容效度、权重的精确数值与工具信效度尚需通过德尔菲咨询、量表研制、层次分析法与学生纵向追踪检验。

关键词:中医药教育;拔尖创新人才;胜任力评价;指标体系;教育评价改革
中图分类号:G642.0 **doi:**10.3969/j.issn.1003-305X.2026.10.712

Design of a multidimensional competency assessment indicator system oriented to “future strategic TCM clinicians”*

YANG Zheng[1], LIU Xi[2], FANG Yini[1], LIANG Xue[1], LIN Jie[1], ZHANG Yan[1], LI Yingshuai[1], WANG Qi[1#]
(1 Wang Qi Academy, National Institute of TCM Constitution and Preventive Medicine, Beijing University of Chinese Medicine, Beijing 100029, China; 2 Office of Academic Affairs, Beijing University of Chinese Medicine, Beijing 100029, China)

Abstract: Objective To translate the four-dimensional competency framework established in the Wang Qi Class curriculum into an assessable multidimensional competency assessment indicator system for cultivating “future strategic traditional Chinese medicine (TCM) clinicians.” **Methods** A cross-sectional questionnaire surveyed 30 experts for an exploratory needs diagnosis to reveal the gap between the perceived importance of competencies and the allocation of promotion weights. Through text analysis and expert-panel argumentation, the indicator system was developed based on the established four-dimensional framework. **Results** Although experts regarded strategic and innovation competency as the foremost reform direction, they assigned the highest weight to grade point average (GPA), with a mean weight of 46.5 points versus 11.6 points for strategy/innovation

杨正,男,博士,讲师
#通信作者:王琦,男,教授,主任医师,博士生导师,中国工程院院士,国医大师,E-mail:wangqi710@126.com
*基金项目:北京中医药大学教育科学研究课题(No.XJZ26036)

(out of 100); among experts highlighting strategic competency, 81.8% still weighted GPA higher. The system covers four dimensions (academic inheritance, thinking training, competency shaping, value guidance), is refined into 16 second-level indicators and six categories of core tools, and specifies an assessment-design pathway for each of the higher-order competencies that are difficult to measure: competency definition—behavioural anchor—assessment tool—scoring rule—decision rule. Three additional targeted design features are proposed: non-compensable threshold gatekeeping of bottom-line values (medical ethics, patient safety); adjustment of promotion weights; and a proposed approach to be piloted "teacher-student comparison of independent pattern differentiation for the same patient", which uses teacher-student agreement on pattern differentiation, supported by agreement on treatment methods, to assess indirectly the tacit, apprenticeship-transmitted competence in pattern differentiation (e.g., integrating and balancing constitution identification, disease diagnosis, and pattern differentiation in a given case). Benchmarked against the five operationally defined attributes of a "strategic TCM clinician" adopted in this study, classical grounding and clinical pattern differentiation are relatively well covered, whereas potential for theoretical originality, strategic awareness and leadership potential show progressively weaker instrumentation, constituting the principal current gaps. **Conclusion** The system offers a to-be-validated assessment-design pathway for challenges such as the difficulty of measuring tacit, apprenticeship-transmitted competence in pattern differentiation and the compensability of bottom-line values by scores; its content validity, weight coefficients, and the reliability and validity of the tools remain to be tested through Delphi consultation, scale development, the analytic hierarchy process and longitudinal tracking.

Keywords: TCM education; top-tier innovative talents; competency assessment; indicator system; educational evaluation reform

《教育强国建设规划纲要(2024—2035 年)》提出"完善拔尖创新人才发现和培养机制""深化教育评价改革"[1],《深化新时代教育评价改革总体方案》要求"坚决改变用分数给学生贴标签的做法"[2]。科学评价拔尖人才的高阶能力,已成为中医药人才培养机制改革的难点。

北京中医药大学(以下简称"本校")"王琦班"由王琦院士领衔,实行本博九年一贯制,培养目标是"有理论产出的中医临床家、有战略思维的中医科学家"[3]。根据这一培养目标、培养方案的四维能力框架和专家论证,本研究对"战略中医临床家"作操作性界定:兼具经典根基、临床辨证、理论原创潜力、战略意识与领军潜质 5 项特征。中医人才的成长靠"读经典、做临床、跟名师",能力高低多在跟师临证的情境中显现,单凭课程考试难以衡量;培养目标与评价手段一旦脱节,重要的能力就容易让位于便于测量的指标。

国际医学教育已形成较成熟的评价框架:胜任力本位医学教育(competency-based medical education,CBME)以能力结果为导向[4],CanMEDS 以 7 个角色界定核心胜任力[5],可信赖专业活动(entrustable professional activities,EPA)把抽象的胜任力转化为可观察、可委托的临床活动[6],里程碑分级追踪能力发展[7],项目化评价聚合多点低利害评价以支撑高利害决策[8],共同构成多源、动态的能力本位评价体系[9]。但它们多来自西医毕业后教育的情境,直接移植难以契合中医经典筑基、师承传承与辨证论治的教育传统。

梳理部分中医药院校公开的培养方案与考核实践可见,国内学业评价仍以课程考试和学业绩点(grade point average,GPA)为主,对战略思维、创新与领导力等高阶能力缺乏系统化评价工具,师承效果也多依赖导师的主观评价。近年来的研究已从长学制院校教育与师承结合[10]、中医师岗位胜任力指标构建[11-13]、拔尖创新人才培养规格现状[14]等方面推进,但多针对已执业的中医师和"岗位达标",较少涉及培养过程中高阶能力的发展性评价,也较少把师承传承单列为一个评价维度。因此,本研究先用 30 名专家的问卷诊断评价取向,再依托王琦班已有的四维框架构建评价指标体系;后续开展德尔菲咨询、量表信效度检验和学生纵向追踪。

1　对象与方法

1.1　总体设计

采用"实证诊断—体系构建"两阶段设计,核心

产出是第2阶段的评价指标体系：第1阶段是横断面专家问卷调研，属于探索性的需求诊断；第2阶段根据其结论，综合多源资料构建指标体系和配套的工具与实施路径。

1.2 调研对象与工具

第1阶段以目的抽样为主，辅以便利抽样，邀请中医临床、科研、教育与教育管理等领域的专家，回收有效问卷30份。专家来自11个省/直辖市18家机构；正高级职称17人、副高级职称5人、中级职称8人，高级职称占73.33%；从业年限均值24.3年(中位数27年，3～42年)。工具是自编的《“王琦班”培养方案调研问卷(专家版)》，共5部分20题，题型含5点李克特量表单选题、多选题(建议最多选3项，超过3项者按实际选择全部计入统计)、100分制权重分配题与开放题，未作信效度检验。

1.3 指标体系构建方法

第2阶段采用多源互证的混合方法，资料来源为文献分析、政策文本[1-2]、王琦班培养方案及配套制度文本，以及2025年8月全国专家论证会与院教学指导委员会专家意见。体系构建分5步推进：维度承接、指标细化、工具配置、权重设定、实施保障；其中指标细化又分建立初始条目池、合并去重、专家论证修订、归类定稿4个步骤。本研究属设计性研究，各级指标的内容效度(拟用2轮德尔菲咨询[15]、计算I-CVI/S-CVI[16-17])和权重的精确数值(拟用层次分析法经一致性检验确定[18])留待后续验证，本文只给出权重的方向性结构。

1.4 数据处理

核心题项由人工录入和程序解析交叉核对，不一致处对照原始问卷逐项确认，异常或缺项由专家本人核实更正。经核校，课程总体评价、衔接度、核心挑战、改革方向与权重配置5题都获得30份完整有效作答(n=30)；权重配置题5项合计100分，各百分比以有效应答人数为分母。分析以描述性统计为主，并对同一专家的挑战/方向选择与权重配置作个体内配对分析；开放题作主题归纳，不作推断性检验。

1.5 伦理说明

本研究是面向专家的意见征询，不涉及患者诊疗信息、生物样本或学生敏感信息。专家在知情同意下自愿具名作答，问卷首页写明所收集信息只用于本项目并保密；回溯核对在去标识化前完成，之后以保密编号存储和分析。

2 结果

2.1 评价取向的专家问卷诊断

2.1.1 现状评价与核心挑战

课程总体评价中，14人(46.67%)选择“亟待改革”或“存在不足”，仅1人选择“非常卓越”；对理论与实践衔接评价为“一般”及以下的专家占33.33%。核心挑战依次为“缺乏系统性的战略思维与创新能力培养模块”(63.33%)、“对学生中医思维的系统性培养不够”(50.00%)和“中西医课程内容比例失衡或结合生硬”(43.33%)。专家的关注点集中于高阶能力培养的系统性缺位，而非具体知识缺失。具体见表1。

表1 当前课程设置面临的核心挑战

核心挑战	选择频次/次	选择率/%
缺乏系统性的战略思维与创新能力培养模块	19	63.33
对学生中医思维的系统性培养不够	15	50.00
中西医课程内容比例失衡或结合生硬	13	43.33
经典理论与现代科技融合不够	10	33.33
课程体系整合度不高，内容交叉重复	9	30.00
实践教学环节薄弱，质与量不足	8	26.67
评价方式侧重知识记忆、忽视能力考核	8	26.67
课程体系固化，更新滞后	7	23.33
教学方法与手段相对传统	5	16.67
人文素养与职业精神培养欠缺	4	13.33

注：本题为多选(建议每人最多选择3项)，部分专家实际选择超过3项；应答30人，选择率以有效应答人数为分母，各项合计大于100%。

2.1.2 改革优先方向

改革优先方向中，“系统性加强战略思维、科研创新能力培养”占60.00%，排在第一位；后序为“优化中西医课程的整合与衔接”(56.67%)、“深化经典学习，强化中医思维根基”(53.33%)、“促进中医与现代科技的深度融合”(43.33%)。深化经典学习与加强战略、科技能力同处前列，二者并不排斥。具体见表2。

2.1.3 重要性认知与权重配置的落差

专家在“重要性认知”上高度关注战略与创新能力；但在晋级/淘汰权重配置(把100分分配给5类指标)上却出现错位：GPA权重均值为46.5分，临床跟师、科研跟师、战略/创新项目、导师与同行评价分别为16.6分、13.9分、11.6分与11.4分；被列为首要挑战与首要改革方向的战略/创新能力分值最低

表 2　课程改革最应优先强化的方向

改革优先方向	选择频次/次	选择率/%
系统性加强战略思维、科研创新能力培养	18	60.00
优化中西医课程的整合与衔接	17	56.67
深化经典学习,强化中医思维根基	16	53.33
促进中医与现代科技(如 AI、大数据)的深度融合	13	43.33
大幅加强临床与行业实践教学	11	36.67
改革评价体系,引入过程性和能力评价	8	26.67
建立高度整合的模块化课程体系	7	23.33
丰富医学人文、职业伦理与沟通课程	3	10.00
推广基于问题教学法等以学生为中心的教学方法	2	6.67

注:本题为多选(建议每人最多选择 3 项),部分专家实际选择超过 3 项;应答 30 人,选择率以有效应答人数为分母,各项合计大于 100%。

(中位数 10 分),56.67% 的专家给 GPA 的权重不低于 50 分。

为避免用群体均值误判个体,采用个体内配对分析。在核心挑战或改革方向题中至少有 1 项突出战略能力且权重配置题有效作答的 22 名专家中,18 名(81.82%)给 GPA 的权重高于战略/创新项目,按改革方向、核心挑战筛选的子组也一致(77.78%,78.95%);权重配置题有效作答的全样本(n=30)中,25 名专家给 GPA 的权重高于战略/创新项目,5 名持平。这里比较的是问卷情境下的权重配置偏向,不是对 GPA 或战略能力价值的直接判定。具体见图 1。

2.1.4　评价体系的设计要求

这一落差指向“重要而难测”的评价工具空白:现有工具还不足以支撑对高阶能力的可信测量;已有研究也指出,胜任力本位评价在实施中面临工具开发与标准统一的挑战[19]。据此提出 3 项设计要求:为战略思维与创新能力提供可量化、可计分的指标;把中医经典筑基、师承传承与辨证论治列为评价的核心内容;使评价随培养阶段递进调整,与九年一贯制学制衔接。

2.2　多元能力评价指标体系的设计

本体系沿一条转化链,拟把高阶能力逐级落实为可观察、可计分的评价内容:能力界定—行为锚点—评价工具—计分规则—决策规则。

2.2.1　指标架构

体系依托王琦班培养方案已定的学术传承、思维训练、能力塑造、价值引领 4 个一级维度,把各维

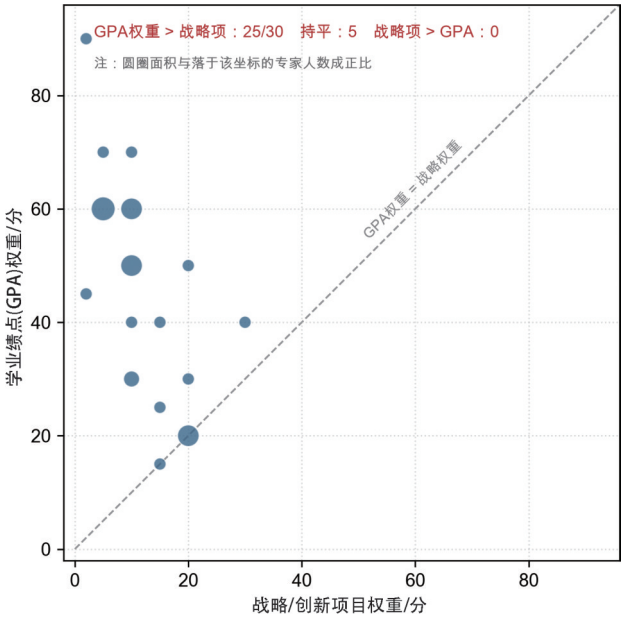


图 1　专家个体晋级权重配置散点

(权重配置题有效全样本,n=30)

注:圆圈面积表示落于同一坐标的专家人数,点落于对角线上方表示该专家给 GPA 的权重高于战略/创新项目。

能力要点细化、重组为 16 个二级指标,并示例列出三级观测点,即转化链中的行为锚点。理论依据是 CBME“以能力为本、结果导向、学习者中心、淡化时间本位”4 项特征[4],并借鉴以可观察活动锚定能力的 EPA[6]、里程碑发展性分级[7]与项目化评价多点聚合[8];同时立足中医“读经典、做临床、跟名师”的成才规律,参考中医原创思维模式[20]、师承重在传思维之道而不止于方药之术[21-22]、名老中医原创思维培养[23]、国医大师成才规律[24]与师承促进职业胜任力的研究[25],使评价维度贴合中医的成才逻辑。具体见表 3。

四维之间存在交叉,因此采用“主要评价意图归属”原则:一项能力关联多个维度时,按最核心的评价意图归入一个维度,配套不同工具,不重复计权。以“辨证”为例,从思维过程(指标 2.1)、临床表现(指标 3.1)、师承传承(指标 1.2)3 个意图分别测量,互不替代。“战略决策思维”(指标 2.3)并非泛指公共政策分析,而是指面向中医药事业发展的研判与规划。

2.2.2　评价工具与计分规则

指标本身不能直接计分,须借助工具生成可供评分的记录。4 个维度共配置 6 类核心评价工具,旨在将师承中难以言传的隐性能力[26]转化为显性证据。具体见表 4。

表3 面向“战略中医临床家”的多元能力评价指标体系

一级维度	二级指标	三级观测点(示例)
学术传承	1.1 经典筑基	经典原文记诵与理解、经典理论的临床运用
	1.2 名师传道	跟诊出勤与质量、跟师日志与作品集完成度
	1.3 临床深化	典型医案整理规范性、同案独立辨证与拟定治法并说明依据
	1.4 传承创新	导师学术思想综述、经验传承论文
思维训练	2.1 经典辨证思维	辨证论治推演(议病识证、议药配方)、整体观与恒动观运用
	2.2 学科交叉思维	中西医知识整合、跨学科问题分析
	2.3 战略决策思维	学术瓶颈与传承创新路径识别、重大疾病与突发公共卫生事件应对研判、政策模拟
	2.4 批判创新思维	文献批判性评价、原创性问题提出
能力塑造	3.1 临床诊疗能力	四诊辨证、理法方药、医患沟通
	3.2 科研创新能力	研究设计、课题申报、科研产出
	3.3 国际对话能力	国际化实践参与、学术交流表现
	3.4 引领协作潜质	团队协作与领导表现、重大问题攻关
价值引领	4.1 家国情怀	服务国家战略的意识、社会实践表现
	4.2 文化自信	中医文化认同与传播
	4.3 医者仁心	医德医风、患者安全意识
	4.4 科学精神	学术诚信、严谨求实

注:一级维度直接采自王琦班培养方案既定的四维能力培养框架;培养方案各维原列能力要点共17条,本研究沿用本校毕业要求“培养要求—二级指标”体例,经合并、拆分与重命名重组为16个二级指标;各级指标的内容效度、行为锚点与权重尚待德尔菲咨询、量表研制与纵向追踪验证。

3项最难测的高阶能力,转化路径分别如下。①隐性辨证能力:行为锚点为学生就同一患者同次临诊的四诊信息独立作出证候判断与治法并说明依据,其记录须在导师给出结论与推理之前提交并锁定,工具为“师生独立临诊比对”,由第三方按统一评分标准盲评;辨证层为主要评分维度,兼评证候判断及其依据的质量,治法层按八法标签与具体治法分项编码,仅单列报告,不计入辨证层评分,也不在指标3.1重复计权;方药由指标3.1另行拟评,本工具不记录、不比较。该工具借鉴在不确定情境下与专家判断相比对的思路,如脚本一致性测验(script concordance test,SCT)[27],但未采用专家参照组与聚合计分,尚不构成同类测验。②战略决策思维:行为锚点为对真实战略议题的研判与规划,工具为战略项目答辩与政策模拟分析报告,按论证质量与可行性计分。③引领协作潜质:行为锚点为团队协作与攻关中的实际表现,工具为多源过程性评价与行为观察记录[28]。相比之下,临床诊疗能力有相对成熟的通用工具可借鉴:如引入迷你临床演练(mini-CEX)与客观结构化临床考核(objective structured clincial examination,OSCE),把四诊辨证、理法方药与医患沟通转化为结构化测评[29-30]。

2.2.3 权重结构与决策规则

转化链的末端是决策规则。表5为拟设的晋级/淘汰决策方向性权重:左列为30名专家问卷配置均值,反映评价惯性下的个体偏好;右列为2025年8月全国专家论证会(10余位校外特邀专家)及院教学指导委员会两次审议形成的集体方向性意见,属于专家咨询,与后续两轮德尔菲不同;两者都不是制度实际执行的权重。

调整方向是:降低GPA的绝对主导,降低导师与同行评价在高利害决策中的权重(其价值底线把关功能另由门槛性指标与多源过程性评价承担),相应提高临床跟师、科研跟师与战略/创新项目的占比;战略/创新项目由问卷均值11.6分调整为拟设15%。表3界定“评价什么能力”,表5界定决策依据何种证据及其权重,两者并非一一对应:同一类证据可支撑多个维度的评价,GPA是各维度的共同基础,因此不单列在某一维度下。

决策规则上,区分“门槛性”与“补偿性”指标:医德医风、学术诚信、患者安全等价值底线属门槛性指标,是必须达标的前置底线,独立于上述权重之外,不能用高绩点或高科研产出抵消;学业绩点、跟师表现、战略/创新项目等属补偿性指标,可在设定范围内综合加权。

2.2.4 实施路径与质量保障

评价指标对接王琦班“分阶段、模块化”安排,形成与“阶段性+通关制”衔接的三阶段递进路径:

表4　四维能力评价工具配置(6类核心工具按维度归并呈现)

一级维度	主要评价工具	工具要点	实施状态
学术传承	中医经典能力等级考核;医案书写评分量表;跟师作品集评估;师生独立临诊比对	经典按"识—记—背—诵—解—用"分级;医案按"四诊—病机—处方—调护"评分;师生就同一患者同次临诊的四诊信息分别记录证候判断(主证型、兼证型、病位与病性)、治法(八法标签并记具体治法)与依据,学生记录于导师结论与推理披露前提交锁定,由第三方盲评比对辨证、治法两层(治法层仅作一致性参考,不计入评分)	经典考核、作品集已纳入培养方案;医案量表、独立临诊比对为拟规范化/拟新增
思维训练	战略项目答辩;政策模拟分析报告;案例辨析	围绕真实战略议题答辩,提交政策模拟分析报告并公开答辩;以经典条文在病案中的辨证逻辑设题	夏季学期战略答辩已纳入培养方案;政策模拟、案例辨析为拟新增
能力塑造	临床行为锚定评价(如mini-CEX);客观结构化临床考核(objective structured clincial examination,OSCE);科研档案	将四诊辨证、理法方药、医患沟通等关键活动转化为可观察、可评分条目;科研以"模拟基金申请书匿名评审"等任务考核	Mini-CEX、OSCE为拟新增/拟改造;科研过程档案已纳入培养方案
价值引领	多源过程性评价;行为观察记录;学习作品集	综合学生自评、导师评价与同伴互评;以行为观察与作品集佐证医德、责任与诚信	导师评价、学习作品集已纳入培养方案;多源评价、同伴互评为拟扩展

注:6类核心工具为中医经典能力等级考核、医案书写评分量表、临床行为锚定评价、战略项目答辩、客观结构化临床考核、多源过程性评价;表中条目为各工具在不同维度下的任务形式或证据载体,同一类工具可对应多项任务,如"师生独立临诊比对"与mini-CEX同属临床行为锚定评价,分别用于师承传承与临床诊疗能力的评价。部分工具已纳入王琦班培养方案,部分为本研究结构化设计的拟用工具,评分细则尚待量表研制与预试。

表5　晋级/淘汰决策权重:专家问卷偏好与论证拟设方案对照

评价项目	问卷权重偏好均值/分	论证拟设权重/%	对应能力指向
GPA	46.5	35	中医与医学基础知识
临床跟师表现	16.6	25	学术传承、临床诊疗能力
科研跟师表现	13.9	20	科研创新能力
战略/创新项目综合表现	11.6	15	战略决策思维、引领协作潜质
导师与同行综合评价	11.4	5	价值引领与综合素养

注:第2列为第一阶段权重配置题有效作答30份的权重均值(GPA中位数50、战略/创新项目中位数10)。第3列属拟设初值,非共识测算结果,其精确数值尚待层次分析法与两轮德尔菲咨询确定。门槛性指标独立于本权重表、不参与加权,故表中"导师与同行综合评价5%"仅指门槛性指标达标后的常规加权部分。

基础阶段(第1~3学年)以学术传承与经典辨证思维为重点,把好"医学综合素质评定"关;提高阶段(第4~5学年)以临床诊疗与科研入门为重点;综合阶段(研究生阶段)以科研创新、战略决策与引领潜质为重点,把好"科学研究综合能力评定"关。新增工具先试点再推广。

质量保障拟从3个层面检验:内容效度经两轮德尔菲咨询[15],并计算I-CVI/S-CVI[16-17]、专家权威系数与肯德尔和谐系数[31];权重以层次分析法构建判断矩阵并检验一致性(CR<0.1)[18];评分者信度以表现性工具的多评分者独立评分计算[28]。另拟按Kane效度框架"评分—概化—外推—后果"4类推断积累效度证据[32]。本体系按项目化评价逻辑[8]聚合多个低利害评价点,由评价委员会作出高利害决策,不以单一得分自动决定。

3　讨论

3.1　认知—配置落差的成因

落差成因可能有三:其一,评价具有导向效应,"考什么"很大程度上决定"教什么、学什么",GPA标准明确、长期沿用,容易形成评价惯性[9];其二,评价主体偏好可量化的指标,开放题中也有专家提出"须先制定具体而公平的考核标准";其三,高阶能力的评价工具缺失,已有成熟工具的能力更容易获得关注与投入[19],缺少可执行依据的能力,即使提高权重也难以落实。GPA是综合性的学业总评指标,其权重高于其他任一单项评价项目,这本身有

合理性；上述落差应理解为权重结构整体上过于倚重传统学业指标，而不是专家忽视战略能力。已有研究提示，评价以理论考试成绩为主时，学生容易陷入死记硬背，削弱临床思维与创新能力[33]。

3.2 体系对“战略中医临床家”培养目标的达成度

“战略中医临床家”由经典根基、临床辨证、理论原创潜力、战略意识与领军潜质5项特征界定。逐项对照本体系的指标、工具与权重，覆盖程度并不均衡。

经典根基与临床辨证的覆盖较充分。前者对应“1.1 经典筑基”，以经典能力等级考核按“识—记—背—诵—解—用”分级计分，已纳入培养方案；后者从思维过程、临床表现与师承传承3个维度分别测量，拟用医案评分量表、师生独立临诊比对、mini-CEX与OSCE支撑，行为锚点相对明确。

理论原创潜力的覆盖居中。该项由“2.4 批判创新思维”与“3.2 科研创新能力”共同承担，工具是模拟基金申请书匿名评审与科研过程档案。局限在于：现有观测点多可由规范科研训练达成，而“理论产出”意义上的原创——即提出新的病机认识或诊疗范式——还缺少专门的锚点；以论文与课题为证据载体，存在把“原创潜力”简化为“科研产出”的风险。

战略意识的覆盖较弱。“2.3 战略决策思维”是唯一直接对应的能力指标，工具是战略项目答辩与政策模拟分析报告，夏季学期战略答辩可作为实施场景；但主要证据载体（战略/创新项目综合表现）拟设权重仅15%，评分细则尚未研制。被专家列为首要改革方向的能力，其测量强度与权重仍与重要性不相称；这说明落差不仅在于专家的问卷配置，也部分留在本体系中。

领军潜质的覆盖最弱。“3.4引领协作潜质”与“4.1 家国情怀”虽已列为二级指标，但观测点主要依赖多源过程性评价与行为观察记录，缺少类似mini-CEX的结构化行为锚点；在九年制培养周期内，领导力表现的观察机会也比较有限，该项当前更接近“记录”而不是“测量”。

在设计层面，本体系已把“重要而难测”的高阶能力落实为可观察指标、配套工具与权重结构，并以门槛性指标守住价值底线。但“是否测得准”还不能断言：5项特征中，越接近中医传统成才路径的特征覆盖越充分，越接近“战略”与“领军”的特征工具化程度越低。这一梯度说明，后续德尔菲咨询与工具研制应优先投向战略意识与领军潜质2项。

3.3 研究贡献与中医特色

已有研究多在宏观层面讨论评价改革，或只提出能力框架的初步设想；本研究则完成了从“框架”到“工具设计”的转化。四维框架取自王琦班培养方案，创新之处在于门槛性/补偿性指标区分、权重调整与拟实施的“师生独立临诊比对”3项设计。门槛性指标使育人与执业底线不被分数抵消，避免“唯分数”倾向以加权求和的形式重现。特色在于：指标与工具主要来自王琦班真实的培养方案、师承制度与专家论证；方法上借鉴CBME、EPA、里程碑与项目化评价[4,6-8]，同时纳入经典等级考核、医案评分、师生独立临诊比对等中医特有方式与师承传统[23-24,34]，并参照“辨体—辨病—辨证”等原创诊疗范式[20,35]。

与国内同类研究相比，中医师岗位胜任力指标体系已较成熟：俞晓东等构建6个一级、19个二级、57个三级指标[11]；袁娜等[12]以“洋葱模型”构建以“大医精诚”为核心的多层次模型；杨振宁等[13]形成6个维度、18个因子、124个特征要素的模型；中医药国际化临床人才评价体系[36]及院前急救医师指标研究[31]也采用多维评价。这些研究多面向已有岗位和“胜任力达标”，本研究则面向九年一贯制的“战略中医临床家”，将评价前移至培养过程。王琦班框架把“学术传承”单列为一级维度，已有模型多将其分散在不同层级[12]；本文沿用这一安排并细化为可评价指标。

四维体系与CanMEDS的多重角色思路相通[5]，但以“学术传承”突出中医经典与师承的根基地位，以“师生独立临诊比对”回应辨证论治“医者意也”、难以言传的特质。权重结构应作为引导证据收集、体现价值导向的参照框架，不能取代专业研判。

3.4 局限与展望

本研究依据一项已完成的探索性专家调研，为需求识别与权重调整提供方向性依据，但构建的指标体系的只是待验证的设计框架；各级指标的内容效度、权重的精确数值与工具信效度均须经过独立检验，方能作为高利害评价工具使用。

第一阶段调研另受多重制约。核心挑战与改革方向2道题为具名多选，“战略思维”“创新”等带有政策上的正面色彩，报告比例可能偏高，“落差”的幅度因此可能部分源于两类题型所承受的社会赞许压力不对称；但个体内配对分析在被试内部作比较，避免了用群体均值误判个体，落差在个体层面方向仍然一致，受影响的主要是幅度与外推范

围。81.8%等配对比例是小样本子组的描述性结果,未作推断检验或区间估计。抽样依托单一项目的专家网络,两阶段证据又都来自中医教育评价领域的专家群体,存在选择偏倚。第二阶段拟设权重来自与第一阶段同质的专家群体,可能带有同样的评价惯性;因此右列权重只作为拟设初值,还需由纳入校外、跨学科评委的两轮德尔菲咨询从外部校准。横断面设计也不支持因果或趋势推断。

“师生独立临诊比对”是本研究新设计且尚无实施与信效度证据的工具,风险需专门控制:患者的实际诊疗始终以带教导师判断为准,学生的证候与治法判断只作教学评价的平行记录,不用于处方或其他诊疗决策;实施前取得患者知情同意并单独中请伦理审查,由不参与诊疗的第三方盲评。该工具拟低频、抽样嵌入常规跟诊,先预设同义词与上下位术语的映射规则,师生一致性同时报告原始一致率与偶然一致性校正指标,并另行检验第三方评分者一致性,初期只作形成性评价;是否进入高利害决策,待信效度验证后再定。由于学生的判断不进入诊疗决策,其在低利害情境下的表现不一定反映真实辨证水平;师生所获四诊信息未必完全一致;两层一致性能否排除表述趋同及导师个人与流派的特异性,并外推为一般辨证能力,尚待多病例、多导师与外部效标验证。

后续研究拟分3步推进:近期在小范围内试点“师生独立临诊比对”,建立评分细则与评分者一致性证据;之后在其他中医院校拔尖创新人才培养项目中检验可迁移性;远期推动评价由选拔性、终结性用途转向“评价—反馈—培养改进”的形成性循环。

4　结论

本研究以本校“王琦班”为例,把既定四维框架转化为多元能力评价指标体系:细化16个二级指标、配置6类核心工具,为难以测量的高阶能力逐项建立“能力界定—行为锚点—评价工具—计分规则—决策规则”转化链,并提出门槛性指标、晋级权重调整与“师生独立临诊比对”3项设计。30位专家的需求诊断显示,专家虽把战略与创新能力列为首要改革方向,权重配置却明显倚重GPA,说明高阶能力评价工具亟待建设。对照“战略中医临床家”5项特征,体系在经典根基与临床辨证上覆盖较充分,在理论原创潜力、战略意识与领军潜质上工具化程度依次递减,后续研制应优先投向后2项。本文的贡献在于体系设计;内容效度、权重的精确数值与工具信效度仍需经德尔菲咨询、量表研制、层次分析法和学生纵向追踪检验。

参考文献:

[1] 中共中央,国务院. 教育强国建设规划纲要(2024—2035年)[EB/OL]. (2025-01-19)[2026-06-20]. https://www.gov.cn/zhengce/202501/content_6999913.htm.

[2] 中共中央,国务院. 深化新时代教育评价改革总体方案[EB/OL]. (2020-10-13)[2026-06-20]. https://www.gov.cn/zhengce/2020-10/13/content_5551032.htm.

[3] 王琦. 中医药大学的“双一流”建设[J]. 北京中医药大学学报,2022,45(9):865-870.

[4] Frank J R, Snell L S, Ten Cate O, et al. Competency-based medical education: theory to practice [J]. Med Teach, 2010,32(8):638-645.

[5] Frank J R, Snell L, Sherbino J. CanMEDS 2015 physician competency framework [M]. Ottawa: Royal College of Physicians and Surgeons of Canada, 2015.

[6] Ten Cate O. Entrustability of professional activities and competency-based training [J]. Med Educ, 2005, 39(12): 1176-1177.

[7] Green M L, Aagaard E M, Caverzagie K J, et al. Charting the road to competence: developmental milestones for internal medicine residency training [J]. J Grad Med Educ, 2009, 1(1): 5-20.

[8] Van Der Vleuten C P M, Schuwirth L W T, Driessen E W, et al. Twelve tips for programmatic assessment [J]. Med Teach, 2015, 37(7): 641-646.

[9] Holmboe E S, Sherbino J, Long D M, et al. The role of assessment in competency-based medical education [J]. Med Teach, 2010, 32(8): 676-682.

[10] 谷晓红,闫永红,林燕,等. 坚持传承创新 促进医教协同:北京中医药大学中医人才培养改革与实践[J]. 中医教育,2016,35(3):7-11.

[11] 俞晓东,沈翠珍,黄文秀. 中医师岗位胜任力指标体系构建研究[J]. 中国医院,2022,26(11):34-37.

[12] 袁娜,焦楠,黄宁宇,等. 中医师岗位胜任力模型构建研究[J]. 中医教育,2022,41(6):37-40.

[13] 杨振宁,周桂桐,门辉华,等. 中医医师岗位胜任力模型构建与实证研究[J]. 世界科学技术-中医药现代化,2023,25(3):852-862.

[14] 黄海鹏,曲铁华. 基于中医学专业人才培养方案的我国中医拔尖创新人才培养规格现状分析[J]. 医学与社会,2023,36(11):138-144.

[15] De Villiers M R, De Villiers P J T, Kent A P. The Delphi technique in health sciences education research [J]. Med Teach, 2005, 27(7): 639-643.

(下转第42页)