

大语言模型训练逻辑启发下的预防医学专业判断能力的培养

周启程*, 屈易萃*, 杨建新, 李红霞, 汤雨潇#, 蔡梦宇#

海军军医大学海军医学系海军营养与食品卫生学教研室, 上海

收稿日期: 2026年8月19日; 录用日期: 2026年9月22日; 发布日期: 2026年9月30日

摘要

生成式人工智能环境下, 预防医学人才培养不仅要解决“学过什么”和“做过什么任务”, 还要回答学生如何在持续反馈中形成稳定的专业判断, 并在陌生、动态和不完全信息情境中保持这种判断。本文借鉴大语言模型开发中的反馈对齐与泛化评价逻辑, 提出预防医学专业判断培养框架。该框架不把反馈理解为一次性纠错, 而是将教师、同伴、实践人员和服务对象等多源信息用于校准学生的问题界定、证据选择、风险判断与行动标准, 形成“初步判断-多源反馈-证据复核-方案修订-再次决策”的循环; 同时通过改变人群、资源、证据和社会条件构造复杂情境迁移任务, 检验学生能否在未见情境中重构方案。本文进一步提出以反馈利用质量、方案修订质量和迁移判断表现评价培养效果, 并以营养与食品卫生学相关任务作为应用示例。该框架旨在推动预防医学教学由“给出答案后的反馈”转向“通过反馈形成判断”, 由熟悉案例表现转向复杂公共卫生情境中的可迁移专业能力。本文将该框架奠基于Kolb的经验学习循环与Vygotsky的最近发展区理论, 阐明反馈修订循环如何把学习过程操作化为可观察、可评分的教学活动, 并明确界定大语言模型类比的适用边界与人类判断中价值、情感维度的不可替代性; 进而提出分阶段、分层次的实施策略: 低年级先行引入教师与同伴的双源反馈, 高年级逐步纳入实践与社区反馈, 并借助在线平台异步收集与管理多源反馈以降低组织成本。

关键词

大语言模型, 预防医学, 反馈, 专业判断, 复杂情境迁移, 经验学习

Cultivation of Professional Judgment in Preventive Medicine Inspired by the Training Logic of Large Language Models

Qicheng Zhou*, Yicui Qu*, Jianxin Yang, Hongxia Li, Yuxiao Tang#, Mengyu Cai#

*共同第一作者。

#共同通讯作者。

文章引用: 周启程, 屈易萃, 杨建新, 李红霞, 汤雨潇, 蔡梦宇. 大语言模型训练逻辑启发下的预防医学专业判断能力的培养[J]. 职业教育发展, 2026, 15(10): 116-124. DOI: 10.12677/ve.2026.1510419

Department of Naval Nutrition and Food Hygiene, Department of Naval Medicine, Naval Medical University, Shanghai

Received: August 19, 2026; accepted: September 22, 2026; published: September 30, 2026

Abstract

In the era of generative artificial intelligence, the training of preventive medicine professionals must address not only what students have learned and which tasks they have performed, but also how they develop stable professional judgment through sustained feedback and maintain it in unfamiliar, dynamic, and information-incomplete situations. Drawing on the logic of feedback alignment and generalization evaluation in large language model development, this paper proposes a framework for cultivating professional judgment in preventive medicine. Instead of treating feedback as one-off error correction, the framework uses multi-source information from teachers, peers, practitioners, and service recipients to calibrate students' problem definition, evidence selection, risk judgment, and action standards, forming a cycle of "initial judgment—multi-source feedback—evidence re-examination—plan revision—repeated decision-making." Complex-situation transfer tasks are constructed by varying population, resources, evidence, and social conditions to test whether students can reconstruct their plans in unseen situations. The framework is theoretically grounded in Kolb's experiential learning cycle and Vygotsky's zone of proximal development: the feedback-revision cycle operationalizes the experiential learning loop into observable and gradable teaching activities, while the progressive addition of feedback sources and the controlled increase of situational complexity operationalize scaffolding and its gradual withdrawal. The paper further delineates the boundary of the language-model analogy, emphasizing that human professional judgment involves values, emotion, and ethical accountability that model optimization cannot capture. To support implementation, a staged strategy is proposed: lower-year students first receive dual-source feedback from teachers and peers, while practice and community feedback are introduced progressively in higher years; online platforms are used to collect and manage multi-source feedback asynchronously so as to reduce organizational costs. Effectiveness is to be evaluated by the quality of feedback utilization, the quality of plan revision, and transfer judgment performance, with nutrition- and food-hygiene-related tasks as illustrative examples. The framework aims to move preventive medicine teaching from feedback given after answers are produced toward judgment formed through feedback, and from performance on familiar cases toward transferable professional capability in complex public health situations.

Keywords

Large Language Model, Preventive Medicine, Feedback, Professional Judgment, Transfer in Complex Situations, Experiential Learning

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

生成式人工智能进入医学教育后, 学生获取建议、比较方案和修订文本的速度明显提高, 但专业判断并不会因获得更多答案而自动形成。对预防医学而言, 很多任务都需要在人群层面处理不确定性: 流行病学调查可能面临数据缺失, 疾病监测可能出现指标冲突, 环境与职业健康风险评估需要在证据不足

时提出下一步行动, 公共卫生干预还要同时考虑资源、接受度、健康公平和风险沟通。WHO 2024 年基本公共卫生职能全球胜任力与学习成果框架强调公共卫生教育与实际职能和学习成果的衔接[1]。

现有公共卫生与预防医学人才培养改革持续强调实践能力与复杂问题解决, 但教学中的反馈仍容易停留在“教师指出错误-学生修改答案”的层面。医学教育研究早已表明, 反馈的作用不仅取决于提供了什么信息, 还取决于学习者如何理解、采纳并把它用于后续表现[2]-[4]。如果学生每次完成任务后只获得分数或标准答案, 就很难观察其专业判断如何发生变化。

大语言模型开发中的反馈对齐提供了一个可供借鉴的工程化参照。经过基础训练后, 模型还需要利用示范、偏好和反馈进一步调整输出行为, 使其更符合既定目标[5]; 与此同时, 训练任务上的良好表现并不能代表面对未见问题时同样可靠, 因此还需通过新任务和变化情境检验泛化能力。人类学习与模型优化机制并不相同, 但其中“通过反馈不断修正专业判断, 并在陌生情境中检验其稳定性和迁移能力”的过程意识, 对预防医学专业判断的培养具有启示。本文据此讨论如何把反馈从课后纠错转变为判断形成机制, 并通过复杂情境迁移评价学生能否真正把专业标准内化并灵活运用。

这一过程意识并非工程学的独创, 而是与成熟的学习理论高度契合。Kolb 的经验学习循环把学习描述为“具体经验-反思性观察-抽象概念化-主动实验”的持续循环, 强调经验必须经过反思并转化为可迁移的概念才能真正促成学习[6]。本框架的“初步判断-多源反馈-证据复核-方案修订-再次决策”正是这一循环的教学化设计: 专业任务提供具体经验, 多源反馈与证据复核构成反思性观察, 修订说明中对判断标准的澄清对应抽象概念化, 再次决策与迁移任务则是主动实验。其推进在于, Kolb 的循环本身并不规定反馈的来源结构与质量要求, 反思结果也常常停留在学生内部而难以观察; 本框架引入科学、方法、实践、社会四类反馈源, 并要求学生把反思成果外显为可评价的修订说明, 使原本内隐的学习循环成为可追踪、可评分的教学证据。Vygotsky 的最近发展区理论则指出, 学习发生在学习者独立解决问题的能力与在更有经验者指导下所能达到的水平之间, 教学的作用在于提供恰当的支持并随能力增长逐步撤除[7]。据此, 多源反馈本质上是一种分布式支架: 教师提供学科标准, 同伴提供可比照的替代思路, 实践人员与社会反馈提供课堂无法复制的现实约束。本框架把支架的撤除操作化为四类情境变化的逐步引入——当人群、资源、证据或社会条件改变时, 原有支架部分失效, 学生必须自行重组判断, 最近发展区的上界因此被有控制地推高。

2. 大语言模型反馈对齐与泛化评价逻辑的教学启示

(一) 反馈应改变判断标准, 而不只是改正答案

在大语言模型开发中, 反馈对齐用于调整模型输出, 使其更符合人类意图和目标要求[5]。将这一思想用于教学时, 不能把“对齐”理解为要求学生服从唯一答案。预防医学很多问题不存在单一最优解, 真正需要校准的是学生判断一项方案是否合理时使用的标准, 例如证据是否可靠、因果解释是否成立、风险是否可接受、方案是否可实施、是否兼顾公平与公共责任。

医学教育中的反馈通常被定义为将学习者实际表现与某一标准进行比较, 并以改善后续表现为目的的信息[2]。但反馈能否发挥作用, 还取决于学习者是否认为反馈可信、是否理解其含义, 以及是否愿意据此采取行动[3][4]。因此, 专业判断培养不能只设计“教师怎么评”, 还要设计“学生如何处理反馈”。

(二) 修订行为是能力形成的关键环节

一次反馈本身并不能证明学习发生。学生真正开始改变专业判断, 往往发生在重新检查证据、比较不同意见、解释为什么接受或拒绝某条建议并再次形成方案的过程中。反馈素养研究强调, 学习者需要能够理解反馈、作出判断并采取行动[8]。在预防医学教学中, 这意味着课程应预留正式的方案修订环节, 而不是把提交作业作为任务终点。

(三) 熟悉任务表现不等于可迁移能力

学生可以在多次类似案例中形成高效率的处理程序,但复杂公共卫生实践经常改变人群、资源、证据和社会环境。适应性专长研究强调,专业人员不仅要在熟悉情境中高效,还要能够面对新问题重新组合既有知识并形成可行方案[9]。因此,课程评价除了观察学生能否把一个任务“做对”,还应检验其能否在关键条件变化后解释哪些判断仍然成立、哪些假设需要调整。

(四) 类比的适用边界与人类判断的独特性

必须明确界定上述类比的适用边界,否则容易把教学简化为参数优化。本文借鉴的仅是两条过程意识:其一,反馈的价值在于进入下一次决策而非终结当前任务;其二,熟悉任务上的表现不能证明能力可以泛化。在三个根本方面,人类专业判断与大语言模型的对齐过程并不相同。第一,优化目标不同。模型对齐追求输出与人类偏好分布的一致,而专业判断追求的是可辩护的标准——学生完全可能且应当在证据支持下偏离多数意见,公共卫生史上许多正确判断在当时都属于少数立场。第二,反馈处理机制不同。模型通过参数更新被动吸收反馈信号,学习者却要判断反馈是否可信、是否与自身经验冲突、是否值得据此改变行为,这一过程受动机、自我效能与身份认同的中介,正是反馈素养研究关注的核心[8]。第三,也是最关键的一点,模型不具备价值观与情感维度。公共卫生判断必然涉及健康公平、弱势群体保护、信任修复与污名化风险等价值权衡:同样的干预方案,在流行病学上有效并不等于在社会意义上正当。风险沟通还要求共情表达——学生需要体会家长的担忧、社区居民的不信任以及一线人员的为难,才能作出既科学又可接受的决策。此外,专业判断伴随真实的伦理责任与后果承担,而模型输出不承担责任。因此,本框架始终把入作为有主体性、有伦理责任和有社会经验的行动者,反馈的作用是帮助学生澄清并捍卫自己的判断标准,而不是要求其服从唯一答案。高质量反馈应当允许学生提出异议,并要求其用证据说明不同选择的合理性。

3. 预防医学专业判断培养中的主要问题

(一) 反馈集中于教师和结果,专业判断过程不易被看见

传统作业和案例讨论中,教师往往在结论形成后给出评价,反馈容易集中在知识错误、格式和结果是否完整。真实公共卫生工作中的方案质量却同时受到方法学、现场条件、目标人群接受度和伦理责任影响。仅由单一教师从学科知识角度反馈,难以呈现专业决策的多维标准。

(二) 学生“收到反馈”,但没有形成可追踪的修订循环

反馈研究反复指出,学习效果受反馈接收、可信度判断和后续行动影响[3][4][10]。现实教学中,很多任务在评分后即告结束,学生可能知道“哪里错”,却没有机会再次完成证据复核和方案调整。这使反馈停留在信息层面,难以转化为判断规则的内化。

(三) 评价情境过于熟悉,难以区分程序记忆与真实迁移

如果期末评价继续使用课堂中高度相似的问题,学生可能凭借熟悉的解题路径获得较好成绩。适应性专长和准备未来学习研究都提示,教育需要让学习者面对变化与不确定性,才能观察其是否能够在新情境中继续学习和调整[9][11]。对预防医学而言,迁移评价尤其需要模拟人群、资源、证据和社会条件的变化。

4. 从反馈修订到复杂情境迁移的专业判断培养框架

(一) 建立“四源反馈”结构

预防医学任务可以至少引入四类反馈。第一是科学反馈,关注概念、研究证据和专业知识是否准确;第二是方法反馈,关注研究设计、数据质量、统计解释和因果推断是否合理;第三是实践反馈,关注时

间、经费、人力、组织流程和可实施性；第四是社会反馈，关注目标人群的接受度、沟通效果、伦理与健康公平。四类反馈不一定方向一致，甚至可能相互冲突，而这种冲突正是训练公共卫生判断的重要材料。四类反馈的来源、核心问题与学生需完成的判断见表 1。

Table 1. Structure of the four feedback sources and the judgments students need to make
表 1. 四源反馈的结构与学生需完成的判断

反馈来源	核心问题	典型提供者	学生需完成的判断
科学反馈	证据和专业结论是否成立	课程教师、领域专家	核查知识与证据，修正事实或机制解释
方法反馈	研究设计与分析是否支持结论	流行病学/统计教师	识别偏倚、混杂、数据限制与替代解释
实践反馈	方案在真实条件下是否可实施	实践导师、基层人员	调整资源配置、流程与执行优先级
社会反馈	方案是否可接受并符合公共责任	目标人群、社区人员	处理沟通、伦理、公平与信任问题

(二) 把“反馈”改造成“判断－修订－再决策”循环

一个专业任务至少应允许学生经历两次决策。第一次方案用于暴露当前判断标准；随后引入多源反馈，要求学生重新检索和评价证据，说明不同反馈之间的冲突，并形成修订方案。学生不仅要提交“修改后答案”，还应回答三个问题：原判断的问题在哪里？哪一条证据或反馈改变了自己的判断？哪些反馈没有采纳，为什么？这种解释使修订过程成为可观察的学习证据。

程序化评价研究强调，应综合多个低风险评价点和反馈信息支持学习，而不是依赖单一高风险考试 [12]-[14]。本框架同样主张将一次任务拆成多个可反馈节点，通过连续证据观察学生判断的发展。

(三) 用“修订质量”而非“修改数量”评价反馈利用

学生改动文字很多，并不意味着专业判断发生了实质变化。修订质量可从四个方面评价：是否识别原方案中的关键问题；是否补充了真正相关的证据；是否改变了不合理的因果链或优先级；是否能够解释修改后的权衡与剩余不确定性。评价的重点应从“有没有听老师的话”转向“能否用专业理由完成更好的第二次决策”。

(四) 设计四类“情境变化”，检验复杂问题迁移

迁移任务不应只是更换病例名称。建议围绕四类关键变化构造新的公共卫生情境：一是人群变化，如儿童转向老年人、城市转向农村；二是资源变化，如经费、人力、检测和时间受到限制；三是证据变化，如新增研究与原结论冲突、数据质量下降或关键指标缺失；四是社会变化，如公众信任下降、媒体关注升高、文化接受度不同或健康公平问题突出。学生需要在原有方案基础上重新判断，而不是重新套用标准流程。

5. 教学应用：以营养与食品卫生学相关任务为场景

(一) 儿童肥胖干预：让不同反馈真正发生冲突

以社区儿童肥胖干预为例，学生首先基于膳食、运动、家庭和学校数据提出初步方案。科学反馈可以关注能量平衡和干预证据，方法反馈检查调查设计与混杂因素，社区实践人员评价方案在学校和家庭中的可执行性，家长或服务对象反馈则揭示接受度和污名化风险。学生需要把这些意见放到同一决策框架中，而不是逐条照单全收。

第二轮任务可改变条件：经费削减、学校无法增加体育课、家长对体重管理产生抵触，或目标人群同时存在营养不足与肥胖。此时学生需要解释原方案哪些部分失效，并重新排序干预优先级。

(二) 食品安全事件：从“证据正确”走向“风险判断可迁移”

食品安全事件更适合训练证据不确定和社会压力下的判断。初始任务可提供污染检测、暴露资料和

毒理学证据, 要求学生形成风险判断与沟通建议; 反馈阶段由不同教师和实践人员指出证据缺口、方法限制和传播风险。迁移任务随后加入“检测资源不足”“媒体报道放大”“不同机构结果不一致”或“公众对官方信息不信任”等变化, 要求学生重新决定哪些信息应优先补充、是否需要采取预防性措施以及如何表达不确定性。

(三) 生成式 AI 作为“挑战性反馈源”, 而不是最终裁判

生成式 AI 可以在教师设定范围内生成备选解释、反例或相互冲突的建议, 用于扩大反馈样本。例如, 模型可以提出一个看似完整但存在因果过度解释的干预建议, 学生需要指出其证据问题并决定是否采纳。AI 也可帮助快速改变人群、资源和舆情条件, 生成迁移任务变式。但专业判断的评价标准必须由课程目标、权威证据和教师共同把关, 不能把模型输出本身当作标准答案。两类示例任务的环节安排与主要评价证据见表 2。

Table 2. Illustrative teaching applications: correspondence of stages between childhood obesity intervention and food safety incidents

表 2. 教学应用示例: 儿童肥胖干预与食品安全事件的环节对应

教学环节	儿童肥胖示例	食品安全示例	主要评价证据
初步判断	提出学校/社区干预初稿	形成初步风险判断与沟通建议	问题界定、证据选择、初始方案
多源反馈	教师、社区人员、家长反馈	毒理、流调、实践与沟通反馈	反馈理解与冲突识别
证据复核与修订	补充证据并重排干预优先级	补充证据并调整风险建议	修订说明、证据变化、取舍理由
复杂情境迁移	资源削减、依从性下降、人群改变	检测受限、舆情升温、证据冲突	条件变化后的再次决策与解释

6. 实施策略与资源考量

上述框架在教学设计上具有完整性, 但落地时必然面对师资、时间与组织协调的现实约束。若不预先规划实施路径, 多源反馈很可能退化为“教师反馈加一次同伴互评”的形式化操作。本章分析主要挑战, 并提出分阶段、分层次的实施策略与资源保障方案。

(一) 实施面临的主要现实挑战

第一, 协调成本高。四源反馈要求实践导师、基层人员乃至服务对象参与, 而这些人并不承担教学任务, 其时间、地点与教学进度难以匹配。第二, 反馈质量不均。同伴反馈可能包含错误的专业判断, 实践人员熟悉现场却不一定了解课程目标与评价标准, 二者若未经引导反而可能误导学生。第三, 评价工作量大。修订质量需要逐份阅读修订说明并判断证据与理由是否成立, 无法像客观题那样批量评分, 大班教学中尤其吃力。第四, 反馈过载风险。低年级学生知识基础薄弱, 同时接收四类相互冲突的反馈容易产生认知超载与挫败感, 反而削弱其采纳反馈的意愿。这四点提示, 框架不应一次性全面铺开, 而应按学习阶段分层引入。

(二) 分阶段、分层次引入反馈源

建议按年级递进安排反馈源的数量与任务复杂度。低年级(一、二年级)先引入“双源反馈”, 即教师反馈与同伴反馈: 教师保障学科标准的准确性, 同伴反馈则成本低、可比性强, 学生在互评中既能发现他人方案的漏洞, 也会因需要解释自己的判断而深化理解; 此阶段任务以结构化问题为主, 目标是建立反馈素养与修订习惯, 而非处理复杂权衡。中年级(三年级)加入方法学教师与实践导师, 形成“三源反馈”, 任务转向整合型与开放型, 重点是训练证据复核与冲突处理能力。高年级(四、五年级及毕业实践)引入社区人员、服务对象与舆情情境, 形成完整四源反馈, 任务以动态迁移为主, 重点是价值权衡与风

险沟通。各阶段的衔接点见表 3。这种分层安排同时回应了最近发展区的要求：反馈源的增加与情境复杂度的提升同步进行，使支架的撤除始终落在学生可承受的范围內。

(三) 利用在线平台异步收集与管理多源反馈

降低组织成本的关键在于把同步、在场的反馈改造为异步、在线的反馈。可依托在线协作平台承载任务提交、反馈批注与版本对比三项功能：实践导师与社区人员以文字或语音留言方式给出反馈，无需到校参加课堂，单次投入通常控制在 15 至 30 分钟，一学期参与 1 至 2 次即可；学生分组提交方案并在平台上留存修订版本，系统自动记录的修改轨迹为修订质量评价提供了客观证据链，教师不必依赖学生自述判断“改了什么”；反馈条目可按科学、方法、实践、社会四类打标签，学生须逐条回应“采纳 - 不采纳 - 存疑”并说明理由，使冲突处理过程结构化、可评分。生成式人工智能可在教师设定范围内生成备选解释与反例作为“挑战性反馈源”，但其输出须经教师审核后方可进入学生视野，以避免错误信息扩散。

(四) 资源保障与可行性安排

实施不必以全课程同步改造为前提。建议先在一门课程的一个单元试点，验证反馈流程与评价量规后再逐步扩展。具体保障包括：建立反馈提供者名单并明确最低参与频次，避免临时协调；修订质量量规控制在四至五个维度(关键问题识别、证据相关性、因果链修正、权衡与不确定性说明)，以兼顾区分度与评分一致性，并安排双评与仲裁；教师培训的重点不是“如何批改错误”，而是“如何提出引发判断冲突的反馈”，可通过案例工作坊与同行评议示范实现；对低年级学生应明确告知反馈数量上限，并预留课堂时间讲解如何处理相互矛盾的意见，防止反馈过载。

Table 3. Staged implementation pathway for multi-source feedback and transfer tasks

表 3. 多源反馈与迁移任务的分阶段实施路径

实施阶段	反馈源结构	任务复杂度与重点培养目标	组织成本
低年级(一、二年级)	教师 + 同伴(双源反馈)	结构化任务为主；培养反馈素养与修订习惯	低，课内即可完成
中年级(三年级)	教师 + 方法学教师 + 实践导师(三源反馈)	整合型与开放型任务；培养证据复核与反馈冲突处理能力	中，需协调实践导师
高年级(四、五年级及毕业实践)	科学、方法、实践、社会四源完整	动态迁移任务为主；培养价值权衡、风险沟通与责任担当	高，需平台与实践基地支持

7. 讨论与结语

从整个能力形成链看，知识基础使学生具备专业判断所需的认知资源，专业任务训练让知识进入真实活动，而反馈与迁移决定这些活动能否进一步沉淀为稳定且可调整的判断方式。本文借鉴大语言模型反馈对齐与泛化评价的价值，主要在于强化这一过程意识：第一次输出不是终点，反馈需要进入下一次决策；熟悉任务上的表现也不是终点，最终还要观察学习者面对变化时能否重新组织知识和行动。

本文提出的“四源反馈 - 修订循环 - 四类情境变化”框架，为预防医学课程提供了一种较具体的教学设计语言。它可以用于检查课程反馈是否过度集中于教师、学生是否真正有机会利用反馈、评价是否只覆盖熟悉任务，也可以帮助课程团队把复杂情境迁移纳入日常教学，而不是等到毕业实践才第一次出现。

从理论定位看，本框架不是提出全新的学习机制，而是把已有理论转化为可执行的课程设计语言。相对 Kolb 的经验学习循环[6]，它解决了“反思环节难以观察、反馈来源缺乏结构”的问题：四类反馈源给出了反思的具体输入，修订说明给出了反思的外显载体。相对 Vygotsky 的最近发展区理论[7]，它解决

了“支架何时撤除、如何撤除”的操作难题：四类情境变化提供了可分级的支架递减路径。相对适应性专长研究，它把“面对新问题重组知识”这一抽象要求落实为具体的情境变量(人群、资源、证据、社会条件)，使迁移评价有明确的构念效度。尚未解决的是反馈源之间的权重问题：当科学反馈与社会反馈冲突时，何种判断才算合理，目前只能依赖教师的个案裁定，未来需要提炼出可教给学生的权衡规则；同时，多源反馈对教师工作量与学生情绪的影响也缺乏实证证据，需要在试点中重点监测。

需要注意的是，反馈越多并不一定越有效。反馈内容过密、来源相互矛盾或超出学生当前知识基础，都可能增加无效负荷。课程设计应根据学习阶段控制反馈数量与复杂度，并明确哪些标准属于科学底线，哪些部分允许根据情境进行权衡。程序化评价同样提醒我们，多个评价点只有在被有意义地整合并促进后续学习时才有价值[12]-[14]。关于反馈效应的元分析同样表明，反馈的效应量在不同研究间差异极大，其效果高度依赖任务难度、反馈指向的层级以及学习者能否将其用于后续表现[15]。

复杂情境迁移也不等于故意把任务变得更难。其目的在于有控制地改变关键条件，观察学生能否识别变化、解释原有假设是否仍成立并调整方案。适应性专长研究强调整理解、变化和发现对于处理新问题的重要性[9] [16] [17]。因此，高年级预防医学课程可以逐步提高情境变化比例，而低年级仍应以稳定基本知识和方法为前提。

上述框架的落地条件已在第六章作了分阶段安排，后续研究可在具体课程中比较单次反馈、教师反馈与多源反馈对方案修订质量的影响，并进一步设置延迟迁移任务，观察学生能否把某一主题中形成的判断标准迁移到新的公共卫生问题。还可建立“修订质量”量规，将证据复核、冲突处理、方案调整和理由解释作为观察指标。通过这类实证研究，预防医学教学有望从“反馈是否提供”进一步走向“反馈是否真正改变判断”，并从“学生是否会做熟悉案例”进一步走向“学生能否在复杂公共卫生情境中持续作出有依据、可解释和负责任的决策”。

参考文献

- [1] World Health Organization (2024) Global Competency and Outcomes Framework for the Essential Public Health Functions. World Health Organization.
- [2] Van De Ridder, J.M.M., Stokking, K.M., McGaghie, W.C. and Ten Cate, O.T.J. (2008) What Is Feedback in Clinical Education? *Medical Education*, **42**, 189-197. <https://doi.org/10.1111/j.1365-2923.2007.02973.x>
- [3] Watling, C. (2014) Cognition, Culture, and Credibility: Deconstructing Feedback in Medical Education. *Perspectives on Medical Education*, **3**, 124-128. <https://doi.org/10.1007/s40037-014-0115-2>
- [4] Bing-You, R., Hayes, V., Varaklis, K., Trowbridge, R., Kemp, H. and McKelvy, D. (2017) Feedback for Learners in Medical Education: What Is Known? A Scoping Review. *Academic Medicine*, **92**, 1346-1354. <https://doi.org/10.1097/acm.0000000000001578>
- [5] Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., et al. (2022) Training Language Models to Follow Instructions with Human Feedback. *Advances in Neural Information Processing Systems* 35, New Orleans, 28 November-9 December 2022, 27730-27744. <https://doi.org/10.52202/068431-2011>
- [6] Kolb, D.A. (1984) *Experiential Learning: Experience as the Source of Learning and Development*. Prentice-Hall.
- [7] Vygotsky, L.S. (1978) *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- [8] Carless, D. and Boud, D. (2018) The Development of Student Feedback Literacy: Enabling Uptake of Feedback. *Assessment & Evaluation in Higher Education*, **43**, 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>
- [9] Mylopoulos, M., Kulasegaram, K. and Woods, N.N. (2018) Developing the Experts We Need: Fostering Adaptive Expertise through Education. *Journal of Evaluation in Clinical Practice*, **24**, 674-677. <https://doi.org/10.1111/jep.12905>
- [10] Watling, C., Driessen, E., van der Vleuten, C.P.M., Vanstone, M. and Lingard, L. (2012) Understanding Responses to Feedback: The Potential and Limitations of Regulatory Focus Theory. *Medical Education*, **46**, 593-603. <https://doi.org/10.1111/j.1365-2923.2012.04209.x>
- [11] Mylopoulos, M. and Woods, N. (2014) Preparing Medical Students for Future Learning Using Basic Science Instruction. *Medical Education*, **48**, 667-673. <https://doi.org/10.1111/medu.12426>
- [12] Schuwirth, L.W.T. and Van der Vleuten, C.P.M. (2011) Programmatic Assessment: From Assessment of Learning to

- Assessment for Learning. *Medical Teacher*, **33**, 478-485. <https://doi.org/10.3109/0142159x.2011.565828>
- [13] van der Vleuten, C.P.M., Schuwirth, L.W.T., Driessen, E.W., Dijkstra, J., Tigelaar, D., Baartman, L.K.J., *et al.* (2012) A Model for Programmatic Assessment Fit for Purpose. *Medical Teacher*, **34**, 205-214. <https://doi.org/10.3109/0142159x.2012.652239>
- [14] Heeneman, S., Oudkerk Pool, A., Schuwirth, L.W.T., van der Vleuten, C.P.M. and Driessen, E.W. (2015) The Impact of Programmatic Assessment on Student Learning: Theory versus Practice. *Medical Education*, **49**, 487-498. <https://doi.org/10.1111/medu.12645>
- [15] Hattie, J. and Timperley, H. (2007) The Power of Feedback. *Review of Educational Research*, **77**, 81-112. <https://doi.org/10.3102/003465430298487>
- [16] Pusic, M.V., Hall, E., Billings, H., Branzetti, J., Hopson, L.R., Regan, L., *et al.* (2022) Educating for Adaptive Expertise: Case Examples along the Medical Education Continuum. *Advances in Health Sciences Education*, **27**, 1383-1400. <https://doi.org/10.1007/s10459-022-10165-z>
- [17] Branzetti, J., Hopson, L.R., Gisondi, M.A. and Regan, L. (2024) Training for Adaptive Expertise: Why, What, and How. *Academic Medicine*, **99**, 121-121. <https://doi.org/10.1097/acm.0000000000005217>