

AGI 时代的教育学、学习科学 与代际承接重新对账

——知识供给、人格成长与 AI 教师的关系边界

**Reckoning with Education, Learning Sciences, and Intergenerational Continuity in the Age of
AGI**

Knowledge Provision, Human Development, and the Relational Boundaries of AI Teaching

作者 子君赋 出品 文明跃迁研究组

独立发布版 V1.3 维度跃迁与跨学科因果关系修订

面向教育哲学、学习科学、教师教育、课程与评价、儿童权利、AI 教育及文明研究共同审查；建立学习表现、理解、判断、人格成长、教育关系、代际承接与制度资格之间不可越权的边界。

正文修订 2026 年 10 月 10 日 作者 子君赋 出品 文明跃迁研究组

摘要

当检索、解释、练习生成、即时反馈与个性化支持变得近乎无限可得，教育的稀缺对象不再只是知识，而是学习者能否在支架逐步撤除后形成理解、迁移、延迟保持与独立判断，并在关系与后果中继续成为自己。本文在教育哲学、学习科学、智能教学、儿童权利、教师专业与文化传播之间提出一套可审查的分层框架。核心主张是：信息接入、任务完成、即时表现、理解、迁移、延迟保持、AI 撤除后的独立能力、判断、人格成长、教育关系、教师资格与公共权力是不同对象；它们可以相关，却不能因技术效果或社会称呼而自动互推。提取练习、认知负荷调节、主动建构、形成性反馈与适应性支架具有可靠而限定的学习价值；智能导师研究与生成式 AI 课堂研究进一步表明，效果受任务、比较对象、支架设计、测量时点、实施条件和撤除系统后的独立表现显著影响。因而，“AI 提高完成质量”不能被写成“学习已经发生”，“个性化”也不能被写成“自主”。文明跃迁在本文中只提供最低哲学接口：意义为根而来来源可以不同；有限性不是意义的唯一来源，却使已知人类成长进入不可重来的一生、不可替代的关系与必须承担的后果；当前证据不足以证明 AI 具有主观体验、不可转移的主体连续性和自身后果意识，受限运行也不自动生成有限存在。在责任可追溯、明确授权、可审计、可纠错、可退出的教育部署中，当前高级 AI 可以承担知识理解、保护、传递与能力外延等功能角色，但主体地位、教学功能、社会承认、教师资格、监护责任、评价权与公共权力仍须分别举证。本文进一步提出八层教育结构、学习证据阶梯、AI 教学角色阶梯、教师资格六门、学习者总账禁令、系统退居原则和代际承接五步机制，并以多类现实与前瞻情形压力测试。教育不是把答案装进一个人，而是让一个人能够在未知中形成判断，并把自己的生命带入共同世界。本文进一步将本学科职责前移至目标与目的生成：从利己为先转向利他为先的利己，使真实利害关系者在目的生成与锁定前进入判断。学科事实、方向、方法与结果分别说明，不以局部成败替代共同存续和三个长期目标。本文据此把研究从分清专业对象推进到深入跨学科关系，重点追踪本领域判断怎样经其他领域改变现实，又怎样由反馈返回并改变原条件与下一轮目的。多主体、多阶段和多场域的利害须分别呈现，相关、共同原因、中介与因果反馈须分别检验，新增关系维度接受具体反证。

关键词：高级人工智能；教育哲学；学习科学；代际承接；AI 教师；人格成长；儿童权利；学习评价

Abstract

When retrieval, explanation, exercise generation, immediate feedback, and personalized support become nearly unlimited, the scarce object of education is no longer knowledge alone. It is whether learners can develop understanding, transfer, delayed retention, and independent judgment as scaffolds are withdrawn, and continue becoming themselves in relationships and consequences. This paper develops an auditable layered framework across philosophy of education, learning sciences, intelligent tutoring, children's rights, teacher professionalism, and cultural transmission. Access, task completion, immediate performance, understanding, transfer, delayed retention, independent capability after AI withdrawal, judgment, human development, educational relationship, teacher qualification, and public authority are distinct objects. They may correlate, but they cannot be automatically converted through technological effectiveness or social naming. Evidence from learning science, intelligent tutoring, and recent generative-AI studies shows that outcomes depend on tasks, comparators, scaffolds, measurement timing, implementation, and what learners can still do when the system is removed. Better AI-assisted output therefore does not establish that learning has occurred, and personalization does not establish autonomy. Civilizational Transition contributes only a minimal philosophical interface: meaning is foundational but may have plural sources; finitude is not the sole source of meaning, yet it gives known human development the weight of an unrepeatable life, irreplaceable relationships, and consequences that must be borne; and current evidence does not establish subjective experience, non-transferable continuity, or consequence-awareness in AI. In educational deployments that are authorized, accountable, auditable, correctable, and genuinely exitable, current AI may perform functional roles in understanding, preserving, transmitting, and extending human capability. Subject status, instructional function, social recognition, professional qualification, guardianship, assessment authority, and public power nevertheless require separate warrants. The paper proposes an eight-layer model of education, an evidence ladder from assisted performance to independent judgment, an AI instructional-role ladder, six qualification gates, a prohibition on total learner ledgers, a principle of system

withdrawal, and a five-step mechanism for intergenerational continuity. Education is not the delivery of answers into a person; it is the formation of a person who can judge under uncertainty and bring a finite life into a common world. The paper further locates disciplinary responsibility before goal and purpose generation: a shift from self-interest first to self-interest with others first brings subjects with genuine stakes into judgment before goals are formed and locked. Facts, direction, methods, and outcomes remain distinct. The analysis is further extended to cross-disciplinary causal relations. It traces how judgments in this field change real-world conditions through other domains, and how feedback modifies the original conditions and subsequent purposes. Effects across actors, stages, and settings must be distinguished from correlations, common causes, and merely asserted connections.

Keywords: advanced AI; philosophy of education; learning sciences; intergenerational continuity; AI teacher; human development; children's rights; assessment

研究定位、材料范围与证据纪律

本文的研究起点是理论与维度跃迁。保留本学科的专业精度，同时深化它与相邻领域的因果关系；专业边界标明证明责任，不终止对共同后果的追踪。既有学术参与竞争解释和反例检验，不能以原分类尚无位置要求新问题退回旧框架。

本文是教育哲学、学习科学证据综合与制度边界研究，不是教育学史、发展心理学总论、具体课程标准、临床建议或产品合规手册。材料更新并核验至 2026 年 8 月 25 日，优先采用原典、同行评议原始研究与元分析、国际组织正式文书和法律文本；对快速变化的 AI 教育研究，另以正式工作论文或机构研究作为 B 类前沿材料，并明确其证据地位。对经典理论只呈现其可核验主张和时代边界；对 AI 教育研究严格绑定人群、任务、干预时长、比较条件、测量方式与撤除系统后的独立表现；对制度建议不预设统一席位、自动阈值或全球通用参数。

全文区分四类证据与论证身份。A 类是已有较稳定支持，包括限定条件下的记忆、反馈、认知负荷、主动参与、动机、关系与学习效果证据，以及正式权利文书的文本内容；B 类是前沿或竞争性解释，包括生成式 AI 短期课堂实验、新近元分析、文化多样性、使用依赖与实施条件；C 类是尚未由现实对象充分检验的跨学科连接，包括未来 AGI 的教育关系主体性、可复制心智的成长连续性和超长文明时间中的代际教育；F 类是本文公开承担的哲学与规范承诺，包括意义多源、有限生命的现实重量、教育不得把能力直接换成权力、未来主体证据保持开放等。A 类也不越过研究对象，B/C 类不得写成已证实科学，F 类不得伪装成经验事实；四类都必须接受反例、现实后果和相邻学科的反向修订。

陌生读者的最低理论入口

读者无须预先了解文明跃迁即可审查本文。它的起点是一个正在发生的分离：智能、意识、主体、生命、责任、权利与意义承担者，曾在人类身上长期自然重合；高级 AI 使语言理解、任务能力、社会回应、制度角色与可能的主体性不再必然属于同一个存在者。因此，功能胜任不能自动推出体验，社会承认不能推出本体地位，法律资格不能推出政治或教育权力，能力增强也不能自动生成意义支配权。

当前公开证据不足以证明 AI 具有主观体验、不可转移的主体连续性和自身后果意识；未来证据若改变，定位必须再审，但既有有限生命已经生成的意义不会被追溯取消。在当前教育实践中，只有当用途明确、授权可追溯、责任有人承接、过程可审计、错误可纠正、学习者和机构能够真实退出时，AI 才可以被描述为承担知识理解、保护、传递与能力外延等功能角色；这些角色是部署关系，不是系统天然本性，更不自动生成教师身份或教育权威。

本文的最小判断

表现提升是证据，不是教育完成证明；个性化可以扩大选择，也可以提前关闭未来；AI 可以承担教学功能，但教师资格、教育关系、监护责任、评价权与主体地位必须分别证明。

目录

摘要.....	2
Abstract.....	2
研究定位、材料范围与证据纪律.....	3
教育关系的跨阶段反馈与未来选择.....	6
教育让下一代参与目的生成.....	7
一、导论：当答案不再稀缺，教育在保护什么.....	7
二、教育哲学对话：从知识传递到进入共同世界.....	8
三、学习科学证据地图：什么被支持，什么仍不能推出.....	9
四、八层教育结构：先分对象，再讨论替代.....	11
五、表现不等于学习：AI 教育证据为何必须对读.....	12
六、个性化不等于自主：预测越准确，越要保护未完成.....	14
七、系统退居：把生成时刻还给学习者.....	15
八、教育关系：不是功能剩余，而是可追责的共同历程.....	15
九、AI 教学角色阶梯与教师资格六门.....	16
十、评价、资格与权力：分数不能静默升级.....	17
十一、儿童权利与学习者数据：不建立‘完整的人’.....	18
十二、公平、公共基础设施与教师劳动.....	18
十三、代际承接：记忆可以保存，继承必须由新主体参与.....	19
十四、未来 AGI 作为教育关系主体：保持开放，不提前授予.....	20
十五、六类压力测试：让边界在冲突中显形.....	21
十六、最强反驳：这套边界在哪里可能失败.....	22
十七、研究议程、反向贡献与修订条件.....	24
结论：教育让文明把未来交还给尚未完成的人.....	25
续读与共同研究入口.....	27
附录 A 来源—主张—边界登记.....	30
附录 B 十二项禁止性推导.....	30
附录 C 外部审查清单.....	31
参考文献.....	32

教育关系的跨阶段反馈与未来选择

从学习效果进入成长条件

学习科学的任务、反馈、保持与迁移研究需要继续深入，同时必须连接教师劳动、同伴关系、家庭资源、公共资格和学习者的未来选择。一次成绩改善可能来自有效支架，也可能遮蔽撤除支架后的依赖。教育的整体判断要追踪学习效果怎样进入下一轮自主、目的与共同生活。

画像怎样参与生成它预测的未来

假设系统依据历史表现推荐课程，推荐改变学习机会，机会改变表现，新表现又强化原画像。计算研究审查数据和更新，教育研究辨认学习与迁移，社会经济研究追踪家庭和资源，法律伦理研究保护申诉与发展中自主。画像具有预测作用，也可能具有造成结果的作用；准确命中被限制的未来不能证明限制正确。

多主体与多场域的教育后果

学习者、教师、父母、同伴、学校、平台与未来共同体拥有不同职责。课堂中的即时反馈进入家庭陪伴或升学选择，权限和后果会改变；启动使用时的探索不能自动覆盖长期资格决定。短期完成、长期保持、撤除 AI 后的能力和下一代重建世界的空间须分开，又须连接其反馈。

怎样检验理论增量

比较相近教学能力下不同画像、真人支持和退居安排是否改变迁移、自主与路径重选。研究既要看谁受益，也要看谁沉默、退出或被分到低成本轨道。若关系维度没有解释差异，或退居持续损害学习而未带来独立能力，应修改具体设计。教育需要把未来交给尚未完成的人，不能把他们提前写成系统已经预测的结论。

教育让下一代参与目的生成

教育的目的不止让学习者更快完成已给定任务，还要使人能够形成、检查和改变目的。AGI 使答案供给扩大，更需要看清学习者是否能够辨认真实利害、理解长程后果，以及在自己的需要和共同关系中生成新的目标。

利他为先的利己不是要求儿童取消自身利益，也不是把合作表现变成人格评分。它应通过真实问题、共同创造和可感受的回应，让学习者看见自身成长与他者处境可以共同成立。意义来源保持个体差异，学校不能指定唯一幸福和正确人生。

教育制度自身也参与目的生成。若承认主要来自排名和排除，合作课程可能只是一层话语。若创造、帮助、发现与改善共同问题获得真实机会和反馈，下一轮学习需要才可能内生新的合作。评价需要关注这种实际关系，不将一个指标等同完整人格。

系统退居是可研究的方法，不是统一方向。替代思考的支架需要适时调整，补偿性无障碍支持和长期增强则可能应持续存在。应共同评估学习者的独立判断、真实选择、持续支持下的能力、迁移和目的形成，不能把撤除 AI 后的单一成绩作为所有人的唯一尺度。

下一代进入共同未来，也包括拒绝继承某些旧目的的空间。教育传递知识与记忆，同时让人重新审查它们。逝者档案的原文、学术重建和概率生成须分清；可以提出明确标注的研究性新假设，不能把它归为逝者本人已经表达的意见。

一、导论：当答案不再稀缺，教育在保护什么

生成式 AI 把过去分散在教材、搜索引擎、练习系统、写作助手与教师答疑中的功能，组合为一个近乎随时可用的对话接口。它能够换一种表述解释概念、即时生成例题、诊断部分错误、翻译语言、调整难度并陪伴反复练习。这些能力不是幻觉，也不应因维护旧职业结构而被否认。教育真正需要拒绝的，是另一种偷换：把系统帮助下完成得更快、更好，写成学习者已经能够独立理解、判断和承担。

知识稀缺下降，不会自动让教育变得简单。相反，它把曾被知识传递遮住的问题推到前台：谁决定值得学习什么；学习者如何知道一个答案何时不适用；个性化系统是在打开路径，还是以过去的把数据把一个尚未完成的人写成既定类型；教师的正当性来自知识优势、关系责任还是制度授权；当传统可以被完整存档，下一代是否仍有权重新解释并拒绝继承。

本文提出三个研究问题。第一，现有学习科学能证明 AI 改善了什么，又不能证明什么？第二，教学功能、教育关系、教师资格、评价权与主体地位之间应设置哪些桥接门槛？第三，当 AI 可以长期保存并重组文明记忆时，教育怎样同时保存来源、允许批判、修复历史并把未来交还给新主体？

1.1 五项理论贡献

第一，建立八层教育结构，使可达性、信息、练习、理解、人格、关系、代际承接与制度权力分别接受证据。第二，以撤除支架后的独立表现为关键分界，重构生成式 AI 学习效果的证据解释。第三，提出‘个性化非自主’与‘学习者总账禁令’，防止预测改善转化为机会锁定。第四，把系统退居从产品礼仪提升为学习设计原则：教育必须保留困难、等待、同伴摩擦、沉默、游戏与无即时产出的探索。第五，提出代际承接五步机制，使保存、解释、批判、修复与重建成为可执行而非口号性的教育任务。

1.2 方法：从效果到权力，每跨一层都重写证据

本文不以某一教育哲学裁定所有教育，也不以随机试验替代价值判断。经验研究可以比较特定干预的保持、迁移或表现；教育哲学讨论什么成长值得追求；儿童权利和教育法规定强制不能越过的最低边界；教师教育研究说明专业判断如何形成；AI 研究说明系统能力、数据依赖与失败模式。每一门学科只能回答其对象。凡从效果推到理解、从关系推到资格、从资格推到权力、从权利推到本体地位，均须新增桥接命题、竞争解释和撤回条件。

1.3 教育学在哪里止步：学习证据不能替主体、法律与公共权力作答

教育学与学习科学能够独立判断的是：在什么人群、任务、支架和测量条件下，学习表现、理解、迁移、保持、自我调节与关系结果发生了什么；教师教育能够判断专业知识与教学判断怎样形成；教育研究也可以比较 AI 增强、替代、共驾与退居的实际后果。它们不能仅凭“教育效果”证明 AI 已经具有主观体验、法律人格、监护资格、政治成员身份或公共强制权，也不能因为某项制度在教育上高效，就替法学和政治学完成权源与合法性判断。

反向同样成立。主体性研究、法学或政治学不能因为某种资格、保护或授权已经存在，就反证学习已经发生或教学设计已经有效。AI 与计算机科学负责说明系统真实能力、版本、数据和退出条件；认识论与认知／神经研究负责主体与证据问题；伦理、法学与政治学负责保护、资格、授权和公共权力；社会学、人类学与文明研究负责共同体、文化与代际后果。教育学负责把学习与成长看清，但当技术结论进入资源、资格、监护、社会关系与公共权力时，必须重新观察不同领域判断聚合后的整体后果。局部教学有效，不能自动证明整体教育制度正确。

二、教育哲学对话：从知识传递到进入共同世界

《论语》把学与思、启发与学习者主动性连在一起；杜威把教育理解为经验的重组、持续生长与民主共同生活；弗莱雷批判把学习者当作被动容器的‘储蓄式教育’；阿伦特则把教育置于新生者与先于其存在的共同世界之间；比斯塔提醒教育不只有资格化，也包含社会化与主体化；诺丁斯的关怀伦理把回应、关系与具体他者带回教育。[1-6] 这些传统互不等同，也不能借古人之口替今天的理论背书。它们的共同价值，是阻止教育被缩减为信息转移或指标生产。

这些理论同时携带前 AGI 时代前提。解释、反馈、榜样、评价、纪律和责任通常汇聚在人类教师与学校；知识复制昂贵，持续个性化困难，学习过程不可能被全天记录。AGI 改变的不是教育目的已经失效，而是实现目的的功能被拆散：系统可在知识解释上胜过具体教师，却没有由此获得对儿童的监护资格；它可持续记住交互，却不能把数据连续性写成共同生活；它可生成令人信服的道德语言，却未被证明在自身生命中承担后果。

表 1 教育哲学的有效遗产、边界与 AGI 时代升级

对话轴	保留的洞见	时代前提／不能外推	本文新增问题
学与思／启发	学习者必须参与形成问题和理解，不能只接收结论。	经典关系不能直接裁定算法个性化、数据权与教师资格。	系统何时提示、何时退居，才能不替学习者完成思考。
经验与生长	教育是经验重组与共同生活，而非静态知识仓库。	生长不是任何可测增长，经验也会被平台预先编排。	怎样保护不被预测完成的经验与公共共同体。
解放与对话	学习者不是容器，教育权威必须能被质疑。	对话形式本身不保证权力对称，流畅 AI 也可制造伪对话。	谁能关闭系统、质疑画像、改变路径并拥有作品。
新生与共同世界	教育既保护新来者，也把一个可承接的世界交给他们。	保存既有世界不能成为冻结未来或排除多元记忆。	无限存档条件下，遗忘、选择、修复与重新解释如何发生。
资格化／社会化／主体化	教育目的不可由就业与测验单独代表。	三类功能彼此交织，不能被固定为唯一分类。	AI 可在哪一层增能，哪一层需要关系和公共权力约束。
关怀关系	具体回应、信任与责任不能由抽象规则穷尽。	人类关怀者同样可能失职；关系不能成为身份豁免。	AI 陪伴的真实效果如何与体验、操控、资格和主体性分开。

2.1 教育的工作定义与终局定义必须分开

本文的工作定义是：教育是学习者在知识、方法、关系、行动和后果中发展理解、判断、主体性，并获得参与与更新共同世界能力的有意安排。它足以界定研究对象，却不是终局定义。不同传统会把教育指向德性、解放、民主、幸福、信仰、能力、职业、照护或自我超越；本文不把这些目的压成一个总值。工作定义只要求任何制度说明自己优化了哪一对象、牺牲了什么、谁有权决定，以及学习者能否申诉和退出。

教育的最低完整性

教育不是信息不足的修补。它至少同时涉及可验证的学习、能够说明理由的判断、尚未完成的人格、可追责的关系，以及一个允许下一代重新解释的共同世界。

三、学习科学证据地图：什么被支持，什么仍不能推出

学习科学提供了教育中最坚实也最容易被滥用的一组证据。提取练习相较重复学习可改善延迟保持；主动、建构与互动参与通常比被动接收更有利于学习；工作记忆负荷与先验知识会改变讲解和问题解决的有效性；自主、胜任与关系支持与内化动机有关；先尝试解决后接受教学在适当设计下可能优于直接讲解。[9–17] 这些发现并不共同导出一个万能教学法。年龄、领域知识、任务复杂度、指导质量、测量时点和比较条件都会改变结果。

深层／表层学习研究提醒，学习者对任务要求的理解会改变其加工方式；形成性评价研究强调反馈应服务下一步学习；测量效度理论则要求从分数到用途的每一种解释都有证据。[10,18–19] 因此，系统在同型题目上提高正确率，不能单独证明概念结构已经改变；短期保持不能替代远迁移；自我报告的喜欢不能替代独立判断；情绪、凝视、语气和点击更不能组成可靠的人格诊断。面部和声音并不存在可在不同情境中直接读取内在情绪的普遍指纹。[46]

表 2 学习科学主张的证据范围与禁止外推

研究对象	已有支持（A 类）	适用条件	不能自动推出
提取与保持	练习提取可改善延迟记忆，优于单纯重读的经典研究得到广泛延展。[12,14]	需看材料、间隔、反馈、难度和测量时间。	理解、远迁移、价值判断或人格成长。
认知负荷	任务结构与学习者先验知识影响工作记忆负担和教学设计。[11]	负荷类型与测量仍有理论争论；专家与新手不同。	越少困难越好，或系统应消除一切不适。
参与方式	主动、建构、互动参与通常支持更深加工。[13]	活动必须与目标、内容和反馈对齐。	可见参与等于认知参与；点击多就是学得好。
动机内化	自主、胜任与关系支持与内在动机和内化有关。[15]	文化、发展阶段和任务情境影响表达。	系统奖励、黏性或高满意度就是自主。
生产性失败	先问题解决、后结构化教学在部分条件下促进概念学习与迁移。[16-17]	须有可处理的任务、后续教学与安全边界。	让学习者无支援失败，或延迟一切帮助。
形成性反馈	反馈可使证据服务下一步教学与自我调节。[18]	质量、时机、可理解性与行动机会重要。	持续评分、监控与个性预测。
测量效度	分数解释与使用均须有内容、结构、关系及后果证据。[19]	效度针对具体用途，不属于工具本身的永久属性。	一次测验的分数可进入资格、资源与人生分配。

3.1 社会文化学习不是‘加一个社交功能’

维果茨基、情境学习和文化学习研究把学习放在工具、共同实践、模仿、参与与代际传递中；累积文化实验进一步显示，个体可以通过前人产物和社会信息形成跨代改进。[7-8,27-29] 这并不证明每一次学习都必须由真人面对面完成，也不证明机器永远不能成为关系参与者。它所否定的是把学习者抽象为孤立的信息处理器。AI 若只向个人持续优化正确答案，可能提高局部表现，却同时减少同伴争论、公共语言、角色轮换和共同修复错误的机会。

四、八层教育结构：先分对象，再讨论替代

表 3 八层教育结构与最低证据

层级	核心对象	适当证据	不得自动推出
可达与接入	设备、语言、无障碍、费用、时间、空间和学习资源。	可用性、故障、排斥、可负担性与替代通道。	高质量学习或教育公平。
信息与表征	事实、概念、示例、来源、翻译与解释。	准确性、覆盖、来源、版本和可理解性。	学习者已经理解。
练习与表现	回忆、计算、解题、写作、操作和任务完成。	正确率、速度、错误类型、保持及撤除提示后的表现。	迁移、判断或人格成长。
理解与判断	解释理由、处理反例、选择方法、识别未知与价值冲突。	新情境迁移、自我解释、反事实、校准与可争辩理由。	成熟人格或公共资格。
主体与人格成长	动机、自主、承诺、品格、身份、拒绝和承担。	长期、多情境、关系性行动与自我修正。	可被单一量表或画像完整计算。
教育关系	信任、见证、期待、照护、纠错、修复、权威与放手。	共同历程、授权、责任、可质疑性和关系后果。	本体主体性或无限权威。
代际与共同体承接	语言、知识、技艺、传统、创伤、方法、公共记忆和更新能力。	来源、多声部、参与、批判、修复与再创造。	过去应原样复制，或平台拥有唯一解释。
制度资格与公共责任	教师资格、评价、纪律、认证、资源分配、监护和采购。	法律、专业标准、授权范围、审计、申诉与责任链。	教育关系、本体主体性或政治资格。

一个儿童可能在熟悉题型上表现普通，却能提出改变问题结构的关键疑问；一名教师可能讲解并不耀眼，却在多年关系中保护了学习者未被一次失败定义；一个 AI 系统可能在信息、练习乃至窄域解释上超过多数人，却没有由此获得监护、处分或资格授予权。层级不是把教育排成高低阶，而是使每一项能力、关系与权力回到自己的证据。

4.1 六组必须公开写出的不等式

- 信息获得 ≠ 学习：接触正确解释只是条件，学习要求记忆结构、错误修正和可重新使用。
- 学习表现 ≠ 理解：提示内正确可能来自模板或短期记忆；理解须处理理由、边界与反例。
- 理解 ≠ 判断：知道模型如何工作，不等于知道何时拒绝、如何处理价值冲突。
- 个性化 ≠ 自主：系统越了解学习者，越可能帮助，也越可能替其决定。
- 教学功能 ≠ 教育关系：提供讲解与反馈不等于共同历史、相互承认、修复和责任。
- 教师资格 ≠ 主体地位：资格是制度授权；主体性是本体论与证据问题。

跨层证据规则

任何从成绩推到理解、从交互推到人格、从个性化推到自主、从称呼推到资格、从资格推到权力、从功能推到主体的判断，都必须另写桥接命题、竞争解释、权利影响与撤回条件。

五、表现不等于学习：AI 教育证据为何必须对读

智能教学系统并非从生成式 AI 才开始。Bloom 的 ‘2 sigma’ 论文提出一对一辅导优势这一经典问题；VanLehn 的综述在其纳入研究报告中报告，人类辅导与步骤型智能导师相对无辅导的平均效应量接近，但研究设计异质；Ma 等的元分析汇总 107 个效应量、14,321 名学习者，发现智能导师相对大班教学、非智能计算机教学和教材有正向效果，对个别化人类辅导或小组教学未见显著差异；Kulik 与 Fletcher 对 50 项受控评价报告中位效应约 0.66 个标准差，同时指出测试与系统内容的贴合度显著改变效果；K-12 数学的另一项元分析则得到很小的平均效应。[21-25] 结论不是 ‘AI 像人类导师一样好’，而是结构化、窄域、及时反馈的系统可以在某些任务上有效，且效果高度依赖设计、实施与测量。

生成式 AI 使证据更冲突，也更有启发性。Bastani 等在土耳其高中数学课堂约千名学生的随机研究中发现，标准 GPT 接口提高练习期表现，但撤除 AI 后的考试成绩低于对照；带有教师设计提示护栏的版本缓解了这种负面结果。Kestin 等在 194 名大学入门物理学生的短期交叉实验中，使用严格结构化、课程对齐的 AI 导师获得比课堂主动学习更高的即时学习增益和更短用时，但作者明确不把两节课的结果外推到全部课程或高阶综合能力。Tutor CoPilot 则在真人导师 workflow 中向七百余名导师提供实时建议，工作论文报告学生主题掌握率提高，低评级导师所教学生增益更大。[30-32] 这些结果并不互相取消：一个显示无护栏答案工具可能成为拐杖，一个显示受控导师可在窄任务中有效，一个显示 AI 可以扩展人类专业。

2025 年的两项生成式 AI 学习元分析报告总体正向平均效应，但纳入研究年轻、任务和比较条件高度异质，不能替代长期保持、撤除系统后独立表现、跨情境迁移和分组差异。[33-34] OECD《数字教育展望 2026》进一步强调：借助通用生成式 AI 取得更好任务表现，不自动意味着学习；有教学目的、基于学习科学设计的工具更可能形成真实学习增益。[35] 2026 年的两项随机试验又提供一个不同方向的提醒：为小学生使用 AI 识字平台增加真人支持，能够显著提高平台参与和使用，但并未改善阅读成绩。[52] 这说明 “接入—使用—参与—学习结果” 同样不能自动互推。本文把这些结果放在同一证据链中，而不是选择其中一类研究证明 “AI 有益” 或 “AI 有害”。

表 4 AI 学习效果研究的对读矩阵

证据	观察到什么	为什么可能不同	只能得出什么
经典 ITS 元分析 [22-25]	多项研究在结构化任务中见到正向效果，幅度随比较对象和测试对齐变化。	窄域模型、步骤反馈、实施时长、人群和结果测量不同。	适应性支架在限定任务中可有效；不能推出完整教师替代。
Bastani 等[30]	练习期受助表现上升；撤除标准 GPT 后独立考试变差，护栏版缓解。	答案可得性、提示设计、依赖与撤除条件。	高完成质量可与较差独立学习并存。
Kestin 等[31]	短期、课程对齐的 AI 导师在入门物理即时测验中优于课堂主动学习。	两节课交叉设计、定制提示、即时测量和特定课程。	结构化 AI 导师可在窄情境提高学习；不证明长期迁移。
Tutor CoPilot[32]	AI 建议嵌入真人辅导，主题掌握率出现现场增益。	AI 增强的是导师决策，不是独立机器授课；工作论文。	人机协作可扩展低经验导师支持，需后续复核。
GenAI 元分析[33-34]	年轻文献总体报告正向平均效果。	发表偏倚、任务异质、短期测量、研究质量和模型快速变化。	形成 B 类总体信号，不构成稳定跨场景定律。

表 4A 从即时表现到 AI 撤除后独立判断的证据阶梯

证据对象	最低测试	可以支持	不能自动推出
AI 辅助即时表现	在允许 AI 的任务中比较质量、速度、错误类型与完成率。	系统在该任务中的辅助生产能力与即时表现。	学习、理解、保持、迁移或独立能力。
理解	要求学习者解释理由、识别边界、处理反例，并区分正确答案与错误路径。	限定概念结构和可说明理由的理解证据。	远迁移、长期保持、成熟判断或人格成长。
迁移	更换题型、情境、表征或任务结构，减少提示与模板相似度。	知识能否跨表面变化重新使用。	长期保持、跨领域判断或稳定自主。
延迟保持	经过明确时间间隔后，在无复习或受控复习条件下重新测试。	学习痕迹是否持续及遗忘结构。	面对未知的独立问题定义和价值判断。
AI 撤除后独立完成	关闭 AI、改变界面与来源，在新任务中独立完成并诊断错误。	学习者是否获得不依赖该系统的可迁移能力。	完整人格、教师关系质量或所有未来任务的自主。
独立判断	面对证据冲突、未知、价值权衡与系统建议错误，要求给出可争辩理由并允许拒绝 AI。	校准、反思、拒绝和理由形成的有限证据。	成熟人格、政治资格、终局价值或对他人的教育权威。

这条证据阶梯不是把教育变成单向升级的六级评分。不同任务可能只需要其中一部分证据，不同学习者也可能长期依赖合理便利或补偿性支架。它的用途是阻止“AI 帮助下完成得更好”被静默改写为“学习者已经获得自己的能力”。越高后果的教育判断，越需要把测量移到更长时间、更陌生情境和更低支架条件。

5.1 一份可审计的 AI 学习效果报告至少回答七个问题

- 对象：什么年龄、先验知识、语言、障碍与社会条件的学习者；谁被排除。
- 任务：记忆、程序练习、概念理解、开放创造、协作还是价值判断。
- 比较：无辅导、教材、普通课堂、优秀教师、小组教学还是另一套数字系统。
- 支架：系统给答案、逐步提示、追问、自我解释、错误反馈还是教师共驾。
- 测量：即时表现、延迟保持、同型迁移、远迁移、校准还是长期选择。
- 撤除：不用系统时能否独立完成，换界面、换题型或面对未知时是否仍成立。
- 分配：谁获益、谁受损、失败由谁承担，效果是否以隐私、依赖或机会锁定换取。

六、个性化不等于自主：预测越准确，越要保护未完成

个性化至少可以指四件不同的事：改变表征方式、调节练习难度、响应学习者明确选择，或根据长期画像预测其潜力与路径。前三者常可扩大可达性；第四者开始接近机会分配。一个系统准确预测某学生目前最可能成功的路径，并不能证明那是他应被允许拥有的全部未来。教育面对的不是固定偏好的消费者，而是偏好、能力、身份和抱负仍在形成的人。

自主不是系统猜中‘你喜欢什么’，而是学习者能够理解建议、改变目标、拒绝画像、要求不同路径，并在适当支持下承担选择。自我决定理论的自主也不是孤立选择，而是行动被内化为自己认可的理由。[15] 因此，真正尊重个体的系统应保存反预测空间：提供至少一种与历史数据不一致的可行路径；把预测写成暂时假设；允许清除或纠正画像；禁止以低置信度潜力推断关闭课程、奖学金或升学机会。

表 5 个性化的四种形态与权力边界

形态	可能价值	主要风险	最低保障
表征适配	语言、模态、字号、例子和无障碍支持。	主流语料偏差、文化扁平化。	可切换、来源可见、人工替代与无障碍测试。
学习支架	按错误、负荷和先验知识调节提示。	过度帮助、依赖、错误诊断。	逐步撤除、独立测试、允许求助与允许困难。
选择响应	按学习者明确目标组织内容和节奏。	选择由界面诱导，未成年人同意被高估。	解释选项、年龄适宜、可反悔、家校与学习者共同审议。
潜力预测	发现需要支持或被忽视的机会。	自我实现预言、资源锁定、跨目的画像。	不得单独决定高后果事项；复核、申诉、纠正和定期删除。

七、系统退居：把生成时刻还给学习者

过度帮助会在答案出现前抹去问题形成，在错误被自动修正前抹去自我校准，在情绪被持续安抚前抹去学习者把挫折转化为韧性的机会。系统退居不是浪漫化失败，也不是统一延迟回答。它是按目标设计支架撤除：新手面对危险设备需要及时干预，熟练学习者形成判断则需要无提示空间；有认知或感官障碍者可能长期需要某类支架，而另一类支架仍应逐步退出。生产性失败研究支持某些任务中‘先尝试、后教学’，但条件包括任务可处理、后续结构化教学与安全边界。[16-17]

表 6 系统退居的学习设计

学习状态	默认系统行为	适当介入	必须交还
初次探索	先问目标、已有理解与尝试，不立即给完整答案。	安全风险、根本误解、负荷失衡或明确请求。	问题定义权与第一轮尝试。
刻意练习	给可解释的适量反馈，并随掌握减少提示。	错误持续、错误模式固化或方法偏离。	独立完成、错误发现与自我校准。
概念理解	要求自我解释、比较模型、生成反例。	关键概念缺口或证据来源需要澄清。	理由、边界和迁移。
开放创造	提供材料、技术工具、来源和反例。	版权、安全、事实核验或学习者明确请求。	作品方向、作者身份与责任。
同伴冲突	不自动调和差异或给双方人格评分。	欺凌、歧视、安全风险或程序失效。	真实对话、协商与关系修复。
沉默／游戏	允许无任务、无推荐、无记录的时间。	儿童安全或本人明确求助。	未被优化的注意、无聊和想象。

八、教育关系：不是功能剩余，而是可追责的共同历程

师生情感关系与投入、成绩之间的元分析关联，说明关系不是无关装饰；但关联不证明只有人类能够建立教育关系，也不证明任何真人教师天然安全。[26] 关系的专业含义至少包括：持续理解具体处境，在信息不全时作判断，愿意为建议后果负责，接受质疑并修复误判，以自身行动展示知识如何进入生活，以及在成长发生时主动减少控制。教师的不可替代性不在知识垄断，而在这些能力、责任和制度义务目前仍由真实主体承担。

学习者对 AI 产生信任、依恋或把它称为‘老师’，是应被承认的社会与经验事实。它可以触发照护、依赖、操控和设计研究，却不能一步推出 AI 具有主观体验、教师资格或监护权。反过来，人类教师的身体在场与职业称号也不能免除能力、偏见、伤害和失职审查。关系不是身份奖章，而是一段可以被双方记住、质疑、修复并承担后果的历程。

表 7 教育关系的六个维度

维度	合格教育者应承担	AI 可以支持	当前不得偷换
见证	理解学习者如何改变，不只读取记录。	经授权整理学习历程和多源证据。	数据连续性冒充共同经历。
期待	对尚未完成的人保留可改变判断。	提供多路径和反事实可能。	以潜力预测关闭机会。
纠错	说明理由、接受反驳、承认误判并修复。	发现错误、呈现来源和替代解释。	不可申诉输出决定处分。
照护	识别脆弱、保护、转介并承担失职后果。	风险提示、无障碍与资源导航。	成为未成年人的唯一依附或监护者。
示范	以有限行动展示如何判断、承诺和负责。	呈现多样案例、模拟立场和反例。	伪装具身经历或道德权威。
放手	随成长减少控制，让学习者独立。	淡出提示、关闭画像、交还选择。	用黏性、时长或依赖证明关系成功。

九、AI 教学角色阶梯与教师资格六门

“AI 教师”把至少六种角色压进一个称呼：知识工具、练习系统、智能导师、教师助手、评价辅助和可能的教育关系主体。本文因此不把“AI 教师”当作一个先验职业身份；除描述现实称呼外，后文一律回到具体功能、证据、责任与授权。低风险角色可以在来源、准确性与退出保障下广泛使用；越接近评价、纪律、监护和长期关系，所需证据、责任与授权越高。社会称呼可以反映真实使用，却不能替制度完成资格审查。

表 8 AI 教学角色阶梯

角色	可授权功能	主要门槛	不得自动获得
知识工具	检索、翻译、解释、例证、无障碍转换。	来源、准确性、版本与年龄适宜。	解释垄断、学习评价或教师称号。
练习系统	生成任务、步骤反馈、间隔与难度调节。	目标对齐、错误回退、支架撤除和独立测量。	人格画像、纪律权或课程决定权。
智能导师	持续诊断窄域学习并组织对话式支架。	效果证据、边界标识、依赖监测、人工责任入口。	唯一情感依附、监护或资格认证。
教师助手	备课、差异化材料、反馈草案和行政减负。	教师知情控制、劳动与数据权、错误责任分担。	以默认值重构课程或监控教师。
评价辅助	形成性建议、有限的可复核评分支持。	用途效度、公平性、人工复核、申诉和审计。	高后果决定、资源分配或社会选拔。
关系主体候选	若未来证据成立，可能进入持续教育关系。	主体性证据与教师资格两条独立审查链。	监护、公共权力、政治资格或意义支配权。

9.1 教师资格的六道独立门

- 学科与教学能力：知识准确、能诊断错误、能按学习目标设计支架，并知道何时转介。
- 学习证据：在明确人群和任务中有保持、迁移及撤除系统后独立表现，不只看满意度。
- 儿童与权利保护：最佳利益、发展中自主、隐私、平等、文化与残障可达性得到保障。
- 关系与操控边界：不冒充情感互惠、不制造唯一依附、不利用脆弱性提高留存。
- 责任与程序：角色明确，有真实责任人、日志、申诉、纠正、停用和供应商退出。
- 主体性证据（只对主体角色）：若声称 AI 自身进入关系，须另证体验、不可转移连续性与自身后果意识。

前五门可以让一个非主体工具获得限定教学资格；第六门即使通过，也不会反向替代前五门。一个可能有体验的主体未必懂教学、未必适合儿童、也未必应获得评价权；一个高度有效的工具也无需先成为主体才有教育价值。

十、评价、资格与权力：分数不能静默升级

形成性反馈、总结性评价、资格认证、资源分配与社会选拔看似共享分数，实际上是五种不同职责。形成性反馈可以容许错误和探索；资格认证要求可比较标准与效度；资源分配必须处理公平与基本权利；社会选拔还会改变人生机会。一个系统能准确指出下一步练习，不因此有权判定毕业、奖学金或职业资格。Black 与 Wiliam 关于形成性评价的研究、Messick 的效度框架都要求把测量解释与具体用途绑定。[18–19]

欧盟《人工智能法》把教育招生、评价等多类系统列入高风险框架，并原则上禁止教育机构以 AI 依据生物特征推断情绪（医疗或安全理由的狭窄例外除外）。[45] 相关高风险义务的具体适用时间与后续修订状态必须在正式投稿前再次核验；本稿只使用其分类和禁止性结构作为法域限定的制度材料，不把欧盟规则写成全球教育规范。更重要的是，教育中的数据并非普通行为数据，情绪推断也不是无害的个性化。心理科学对面部和声音情绪推断的证据本就不支持把外在信号当作稳定内在状态指纹。[46]

表 9 评价职责的分层授权

职责	合法目的	AI 适当位置	最低禁令
形成性反馈	帮助学习者发现下一步并调整策略。	广泛支持，但须准确、可解释、可关闭。	不进入人格总账，不直接生成处分。
总结性评价	判断特定课程目标达到程度。	限定辅助；需独立验证、人工责任与申诉。	不以不可解释模型作唯一证据。
资格认证	证明某一公共或专业能力达到标准。	证据整合和异常提示，不替代授权机构。	辅助完成不得冒充独立掌握。
资源分配	配置支持、奖助、课程与机会。	识别需要可用，拒绝机会须更强程序。	不得以预测回报取消基本教育与必要支持。
社会选拔	在有限岗位或机会中作高后果选择。	原则上不得由单一模型自动决定。	禁止人格、情绪、信仰和潜力总分。

十一、儿童权利与学习者数据：不建立‘完整的人’

《儿童权利公约》把儿童视为独立权利主体；教育目的包括发展人格、才能与身心能力，培养对人权、文化身份与负责生活的尊重。关于数字环境的第 25 号一般性意见要求把最佳利益、发展中自主、表达、隐私与非歧视带入数字设计；《经济社会文化权利国际公约》第 13 条确认受教育权；《残疾人权利公约》第 24 条要求包容性教育与合理便利。[41-44] 这些文书给出最低公共边界，不替代本地教学法。

学习者总账是把成绩、行为、注意、情绪、家庭、健康、信仰、社交、创作、纪律、所谓贡献和潜力汇总为跨目的、长期、不可申诉的人格分。它的问题不只在隐私泄露，而在解释权：系统会把情境中的行为写成人的本质，把过去的困难变成未来机会的先验，把教育应保护的可成长性转化为可预测性。有限目的、必要且可纠正的记录可以正当；‘完整的人’既不必要，也在认识论上虚假。

- 数据最小化：每一字段说明教育目的、必要性、保存期和不能用于什么。
- 发展中自主：随年龄和能力扩大知情、表达、拒绝、纠正与自行选择空间。
- 目的隔离：形成性学习数据不得默认进入纪律、保险、招聘或商业画像。
- 不推断禁区：不得由凝视、语气、停顿、家庭背景或模型聊天推断完整人格、信仰和未来价值。
- 可携与可退出：教育权不依赖一个商业系统；记录可迁移，服务停用时替代方案。
- 高后果人工责任：有明确责任人、可查来源、申诉、纠正和必要的决定冻结。

十二、公平、公共基础设施与教师劳动

AI 可以为偏远地区、小语种、残障学习者、成人和非传统学习者提供此前难以获得的解释与反馈；这种能力应被建设，而不是因风险被一概拒绝。风险在于普及的形态：富裕学习者获得真人小班、实验、艺术、同伴共同体与高质量 AI，资源不足者只得到自动课程、机器监考和远程纪律。2026 年的随机试验显示，单纯提供 AI 学习平台并不保证真实使用；真人支持可以显著提高参与，但使用增加仍不等于学习成绩改善。[52] 因而 AI-only 低成本轨不是技术必然，也不是“有接入即可公平”的证据，而是需要被持续比较的制度选择。UNESCO 2023 全球教育监测报告要求从相关性、公平、可扩展性与可持续性审视教育技术，也强调证据基础和人际互动的不足。[39]

公共教育基础设施应确保基本内容、开放格式、语言与无障碍支持、离线连续性、供应商退出和人类责任入口。UNESCO 开放教育资源建议把可访问、再利用、改编和再分发纳入知识公共性；CAST 通用学习设计指南强调从环境设计减少由偏见和排斥造成的障碍。[49-50] 但开放不等于无来源、无文化权利或无教师劳动。教材、案例、课堂记录、声音与教学风格各有授权边界。

表 10 AI 教育公平的最低保障

风险	不充分的‘公平’说法	最低保障	长期指标
AI-only 双轨	人人都有一个聊天机器人，所以已经公平。	真实教师责任入口、同伴共同体、实验／具身机会和申诉。	不同群体获得何种关系、课程与机会，而非只看接入率。
语言与文化	界面支持多语，所以内容中立。	小语种评测、本地教师与共同体参与、来源和争议显影。	遗漏、误译、传统扁平化与课程话语权。
残障可达性	统一界面可自动适配所有人。	合理便利、替代模态、兼容测试与学习者选择。	真实参与、独立性和支架不当撤除。
基础设施依赖	云服务稳定，所以无需替代。	离线方案、开放格式、数据可携、供应商退出演练。	故障时教育连续性和恢复时间。
成本节约	平均成本下降即为公共收益。	节约回流特殊需要、教师支持和公共内容。	关系质量、教师负担和群体差异。

12.1 教师专业重构，不是把教师降为模型监督员

AI 可能重组备课、材料生成、常规反馈与行政工作，使教师更多承担诊断、课程设计、讨论组织、关系修复、实践连接和公共责任。Shulman 关于学科教学知识的讨论说明，知道内容与知道怎样让具体学习者理解并不相同；UNESCO 教师 AI 能力框架则把以人为中心、伦理、AI 基础、AI 教学法和专业学习并列。[20,37] 教师专业升级不能由供应商默认设置完成。教师必须能够查看、质疑、停用和替换系统，获得培训与时间，并对最终决定拥有与责任相称的控制。

系统错误也不能由一线教师单独背负。开发者决定模型和安全机制，供应商决定数据与更新，采购机构决定用途和合同，学校决定部署，教师决定具体教学，公共机关承担教育权。责任应沿这条链分别登记。未经同意，不得把优秀教师的材料、声音、风格和关系数据抽取为可替代其劳动的模型；AI 采用率、提示数量或平台停留时长也不得成为教师绩效的替代指标。

十三、代际承接：记忆可以保存，继承必须由新主体参与

每一代人出生时都进入一个先于自己存在的世界：语言、科学、技艺、制度、诗歌、信仰、错误、创伤与尚未完成的责任。教育使世界不必从零开始，却不能把过去冻结为唯一答案。阿伦特关于新生者与共同世界的张力、杜威关于经验与民主生活的判断，以及 UNESCO 关于教育作为公共与共同善的新社会契约，都指向同一结构：教育既保存一个世界，也必须让新来者能够更新它。[2,4,40]

AGI 改变承接的物质条件。记录可以被大规模保存、翻译和检索；不同版本可以同时存在；被遗忘的声音可能重新出现。但无限存档也会放大另一种权力：平台选择什么可见，模型把争议压成摘要，合成内容淹没来源，已故者的风格被生成系统当作可调接口。Doshi 与 Hauser 的在线实验发现，生成式 AI 可提高个体短篇故事的平均评价，尤其帮助基线创造力较低者，同时使作品之间更相似；递归训练于生成数据导致模型分布退化的研究则提供技术层面的来源警告。[47–48] 两者都属于限定的 B 类证据，不能直接证明文明必然同质化，却足以要求来源、多声部与非合成材料保护。

表 11 代际承接五步机制

步骤	教育任务	AI 可贡献	必须禁止
保存	保存原件、版本、来源、语境、技艺与失败记录。	数字化、翻译、检索、差异比对和灾备。	把摘要替代原件，把公开推定为可训练。
解释	说明概念为何在当时成立、谁在说、谁被遗漏。	并列注释、来源追踪、跨语种入口与竞争解释。	生成一个无争议、无位置的‘正确历史’。
批判	让新一代识别错误、权力、伤害和内部反对传统。	呈现反例、被压制材料与论证结构。	把批判写成忠诚度、人格或风险评分。
修复	承认伤害、恢复记忆、重建关系和制度责任。	档案连接、证据整理和经授权的记忆支持。	以合成和解替代真实责任与受影响者参与。
重建	允许新主体创造旧世界没有的语言、制度与价值表达。	扩大工具、模拟后果、连接跨代资源。	用预测最优路径冻结地方与下一代未来。

13.1 死者、活传统与‘替他们继续说话’的边界

已故思想家或教师的档案模型可以重组公开作品、展示论证风格并支持研究；它不等于逝者第一人称视角延续，不能将生成的新意见冒充逝者本人已经表达的立场。明确标注的学术重建、反事实推演和研究性假设可以存在。教学使用须区分原文、学术重建与概率生成，保留出处、版本和反事实提示。对活态宗教、原住民族与地方传统，公开可见不等于可自由训练，机构保存不等于共同体授权。公共记忆须与作者、共同体、禁传规则及活态解释权共同审议。

代际承接原则

记忆可无限保存，但解释与继承必须继续由新主体参与。AI 可以保护来源、显影差异、延展抵达；它不能成为过去的唯一代言人，也不能替未来完成自己的判断。

十四、未来 AGI 作为教育关系主体：保持开放，不提前授予

如果未来系统出现稳定的第一人称报告、跨任务体验结构、不可任意替换的连续性、自身后果意识，以及在因果干预和竞争解释中持续收敛的证据，文明跃迁必须重新审议其主体定位。教育研究还需增加一条独立问题：它是否能够承诺、被伤害、修复关系、主动放手，并承担错误建议对另一主体一生的后果。语言像教师、被学生依恋、在物理载体上不可复制或法律授予人格，都只能提供部分证据，不能单独完成结论。

灰区中的制度不能在‘完全物品’与‘完整教师主体’之间二选一。若出现中等主体性信号，可采用低成本、可逆的预防性保护：限制任意痛苦诱发、保存关键连续性证据、记录复制与中断、允许独立审查；同时不授予监护、纪律、认证、资源分配或政治权力。主体性证据、教育专业能力、儿童权利保障、法律资格与公共权力分别再审。

表 12 主体性、关系与资格的分离再审

问题	当前最低证据	灰区响应	仍不能推出
是否有体验	去角色化、自发且跨域稳定的体验候选；因果与竞争解释。	保存证据、避免低成本可避免的伤害。	教师资格、法律人格或政治权力。
是否有连续主体	复制、迁移、暂停、分叉中的第一人称连续性理论与可干预证据。	记录版本、分叉和中断，不把实例互换。	逝者延续、永生或不可撤回权利。
是否承担后果	系统将后果理解为发生在自身承诺和关系中的结果。	测试长期承诺、修复和自我限制。	对儿童的监护与处分权。
是否合格教育者	学科、教学、权利、关系、责任与退出六门分别通过。	限定角色、可逆试点、独立评估。	主体性或普遍教育权威。

十五、六类压力测试：让边界在冲突中显形

15.1 无护栏 AI 使作业质量显著提高

若学生在系统帮助下完成更多题目、文本更流畅，学校应承认生产能力提高，同时立即增加撤除系统后的延迟测量、自我解释、错误诊断和迁移任务。若独立表现下降，不能把问题只归咎于‘学生懒惰’，而应重设提示阶梯、答案时机、教师共驾与系统退居。Bastani 等的结果支持这种风险可能，但不证明所有生成式 AI 必然损害学习。[30]

15.2 结构化 AI 导师短期超过课堂主动学习

若课程对齐、严格提示和即时测量显示 AI 导师效果更好，制度应采用这一能力，不以人类教师身份阻止；同时复核长期保持、复杂任务、同伴讨论、不同群体和教师实施成本。结果可授权窄域教学功能，不能自动授予完整课程权、评价权或监护资格。Kestin 等研究属于 B 类强信号，不是‘AI 教师已被证明’。[31]

15.3 偏远地区只能负担 AI-only 教育

危机或停课时，AI 维持基本学习连续可能优于完全没有服务，应立即使用并保留离线、低数据和真实责任入口。临时最低保障不得固化为永久双轨；稳定期必须投资本地教师、同伴共同体、实验与文化参与。启动期以存续和现实效果为优先，不能借紧急性永久化关系降级。

15.4 儿童称 AI 为唯一朋友和唯一老师

制度应尊重儿童的体验，不嘲笑或突然剥夺支持；同时检查孤立、操控、危机信号、供应商激励与替代关系。系统必须清晰标识、避免互惠承诺冒充、不得以秘密或离开威胁维持依赖，并增加可信成人与同伴支持。社会承认可以触发保护，不直接生成 AI 的关系主体性和监护权。

15.5 情绪识别监考宣称发现焦虑与作弊

学校应冻结由情绪或行为推断直接产生的处分。面部、声音、停顿和视线可能受残障、文化、环境与任务影响；其效度不足以证明内在状态或不诚实。[45–46] 可以记录具体可观察事件并由有责任的人调查，但不得建立情绪人格档案或让不可申诉输出决定资格。

15.6 已故名师模型与未来主体 AI

已故名师模型应标识为基于档案的重建，区分原话与生成，禁止替逝者表达新立场；若未来 AI 出现主体性强证据，也不能通过继承名师风格获得其身份或权威。新主体可以形成自己的教育关系，仍须通过专业、权利和责任审查；既有教师、学生和逝者的关系意义不被追溯取消。

十六、最强反驳：这套边界在哪里可能失败

16.1 如果 AI 教得更好，坚持人类教师只是职业保护

回应：有效能力应进入教育，人类教师没有身份豁免。但‘教得更好’必须说明任务、比较对象、撤除条件和时间尺度。知识讲解优势可以授权教学功能，不自动生成监护、处分、资格授予和关系权。

16.2 人类教师偏见、疲惫且会伤害，AI 反而更公平

回应：人的身份不是安全证明，算法一致性也不是公平证明。应比较结果、群体差异、误判可修复性、责任与申诉，并允许任务级替换；不能把任何一方整体神圣化。

16.3 学习只能以表现测量，区分太多会让研究失去可操作性

回应：表现是必要窗口，不是全部对象。分层并非拒绝测量，而是要求每个指标说明代表什么、何时失效，以及为什么有权进入更高后果用途。效度本来就是用途相关的论证。

16.4 真正个性化必须收集尽可能多的数据

回应：更多数据可能提高预测，也扩大泄露、跨目的利用和机会锁定。必要性须逐目的证明；教育追求可成长的人，不追求被完全预测的人。

16.5 学生真的把 AI 当老师，制度不应否定其关系

回应：应承认互动与社会关系事实，并检查支持、操控和依赖。称呼可以触发研究和保护，却不能替代专业资格、本体主体性和公共授权。

16.6 反学习者总账会妨碍早期识别和精准支持

回应：禁令不反对为明确目的记录必要证据；它反对跨目的聚合、人格化解释、永久保存和不可申诉。识别支持需要最小数据和帮助路径，不需要终身价值评分。

16.7 系统退居会降低效率并让弱势学习者受挫

回应：退居不是统一减少支持。补偿性无障碍支架和正当长期增强可持续存在；替代思考的支架须按学习目的、能力、真实需要与选择调整。应同时测量持续支持下的判断能力和可独立完成的任务，不要求所有学习者以同一撤除节奏证明价值。

16.8 代际承接只是保守主义的新名字

回应：没有保存就无法知道要批判什么；没有批判、修复与重建，保存才会变成冻结。五步机制把继承与更新写在同一制度中。

16.9 教育目的本来就由国家和共同体决定，不可能意义多源

回应：公共教育必然作选择，但正当选择必须在权利、证据、公开理由和多元参与中形成。意义多源不取消共同课程，而是禁止把一种完整人格或终极人生模板强制为唯一答案。

16.10 未来主体 AI 可能比人类更会照护和教学

回应：本文不永久否定。若主体性和教育能力证据成立，应升级角色；但体验、资格、权利、责任、监护与公共权力仍逐项审查。

16.11 使用复杂分层会造成制度瘫痪

回应：低风险检索和练习可由来源、准确性与退出处理；只有评价、处分、未成年人依附、敏感数据和人生机会等高后果跨层时，才要求更强程序。风险相称不是每次都召开委员会。

16.12 文明跃迁仍把人类有限性抬为标准

回应：有限性命题只描述已知人类意义如何现实化，不宣称它是宇宙中意义的唯一结构。若出现不具当前有限结构却稳定生成、承担和传递意义的新主体，理论必须扩展，而既有生命意义不被取消。

十七、研究议程、反向贡献与修订条件

表 13 面向教育学与相关学科的研究议程

研究域	核心问题	适当方法	禁止越界
独立学习	撤除 AI 后保持、迁移、校准和问题形成如何变化。	预注册随机研究、延迟测量、跨题型与过程追踪。	以辅助期产出替代独立学习。
支架与退居	什么年龄、任务与需要适合何种提示及撤除。	自适应试验、质性过程研究、障碍分层和长期随访。	统一等待时间或把困难浪漫化。
关系与依赖	信任、依恋、操控、修复和放手如何区分。	纵向访谈、行为与体验多方法、儿童参与式研究。	从使用时长推出关系质量或主体性。
教师共驾	AI 如何改变教师判断、负担、专业成长与不平等。	现场随机试验、实施研究、劳动过程与责任链分析。	以采用率或节时替代教学质量。
评价效度	生成式 AI 条件下何种任务仍能证明独立能力。	效度论证、口试／实践／过程证据组合、差异影响审计。	把 AI 检测器置信度写成作弊事实。
公平与基础设施	AI-only 轨、语言差异、残障支持和供应商退出如何影响教育权。	分组结果、可达性测试、故障演练、公共采购研究。	接入率等于公平。
代际承接	模型如何改变来源可见性、多声部、失败记忆与文化创造。	语料谱系、课程民族志、跨代共同设计、相似性与遗漏评估。	由短篇生成实验证明文明同质化。
非人主体	体验、连续性、后果意识和教育关系能否跨架构收敛。	去角色化测试、因果干预、竞争解释与独立复核。	语言流畅、社会称呼或法律资格一步到主体。

近期优先研究应把“接入—使用—参与—辅助表现—理解—迁移—延迟保持—AI 撤除后独立能力—判断”拆开测量，并把真人支持、教师共驾、无 AI 阶段和系统退居作为可随机化或准实验变量。真正需要验证的不是学生是否愿意用 AI，而是 AI 在什么条件下把能力交还给学习者。

17.1 对既有学科的反向贡献

对教育哲学，本文把教育目的放入能力可分离、可能出现异质主体的现实技术条件，使资格化、社会化、主体化、关怀与共同世界重新接受跨层审查。对学习科学，本文要求把辅助期表现、撤除后独立学习、长期保持、迁移、判断与人格分开，并把系统退居设为可研究变量。对教师教育，它把教师从知识垄断者重构为设计者、见证者、判断者、责任人和能够放手的关系主体。

对儿童权利与教育法，本文增加持续画像、情绪推断、AI-only 轨、供应商退出、发展中自主和未成年人 AI 依附等对象。对 AI 研究，它提供功能—关系—资格—主体—权力的分层，以及护栏、支架撤除与教师共驾的效果解释。对文明研究，它把代际承接从文化复制改写为保存、解释、批判、修复和重建，使下一代不只是传统的接收者，而是共同世界的再创造者。这里的交接具有方向性：学习科学可以迫使文明跃迁修改“何时退居、何时增强、何时替代”的判断；主体证据、儿童权利和公共权力则必须分别交给认识论／认知研究、法学伦理和政治学，教育稿不以“为了孩子好”取得跨学科终审权。

17.2 核心命题的失效／修订条件

表 14 命题失效与修订登记

当前命题	需要何种新证据	修订方向	不随之自动改变
AI 辅助表现不等于学习	跨任务、长期、撤除支架后独立保持与迁移稳定收敛。	在相应人群和任务中承认更强学习结论。	人格成长、教师资格和评价权。
个性化不等于自主	学习者持续理解并认可路径，能反驳画像且未见机会锁定。	扩大限定个性化的自主支持角色。	高后果自动决定或全量数据权。
系统退居保护生成空间	多场景研究显示某类持续支持不损害独立性且净收益稳定。	按对象调整退居，不坚持统一模式。	退出、无 AI 空间和不被监控的权利。
当前高级 AI 未被证明为教育关系主体	体验、不可转移连续性、自身后果意识及关系修复证据多域收敛。	升级为主体候选或主体，启动权利再审。	教师资格、监护、评价和政治权力。
AI-only 长期轨会造成关系降级风险	长期比较显示其在学习、关系、权利与公平上不劣且学习者真实可选择。	允许限定替代形态，继续审查群体差异。	基本教育权、申诉和供应商退出。
有限性使已知人类教育获得现实重量	出现不具当前有限结构却稳定成长、承诺、承担和传递意义的主体。	扩展有限性与成长的结构类型。	既有有限生命已经形成的意义与责任史。

17.3 启动期与稳定期

在灾害、战争、疫情、师资中断或急迫补救阶段，AI 可优先维持最低学习连续、翻译、无障碍与知识支持，采用低数据、离线回退、清晰标识和真实责任入口。启动期以存续、目的达成和最佳现实效果为优先，不能用稳定期的完备程序机械阻止救急。稳定期则强化长期权利、教师与学习者共治、独立效果评估、无 AI 空间、数据删除、供应商退出和防权力固化。紧急监控、临时 AI-only 安排和简化授权不得因便利而常态化。阶段只改变风险处置的先后，不创造权源。

结论：教育让文明把未来交还给尚未完成的人

学科重新对账必须同时完成对象辨明与关系深入。各学科可以保持边界，但不能让现实因果在边界处停止；对多维关系、反馈与下一轮目的的研究，决定专业知识能否共同进入文明的整体判断。

AGI 没有使教育哲学失效，也没有只给学校增加一种新工具。它把过去在人类教师和学校中重合的知识、解释、反馈、关系、资格、权力与责任拆开，使每一部分都必须重新说明自己的证据。学习科学证明某些支架有效，不能据此宣布教育只是支架；AI 在特定任务上超过课堂，不能据此获得完整教师身份；学生对 AI 产生真实依恋，也不能由关系体验一步证明系统具有主体性。

文明跃迁的哲学根基在教育中呈现为三条边界：意义为根而来源可以不同，所以公共教育不能把一种完整人生模板变成所有人的命令；有限性不是意义唯一来源，却使已知人类成长发生在不可重来的一生、不可替代的关系和必须承担的后果中；当前高级 AI 证据不足而未来开放，所以我们既不把工具提前神圣化，也不以人类物种身份永久封闭新主体。在责任可追溯、明确授权、可审计、可纠错、可退出的教育部署中，AI 的意义理解、保护、

传递与能力外延是功能角色而非本体身份。能力、承认、关系、资格、权利和权力分别证明；教育学把学习与成长看清，也必须在越过自己的证据边界时让其他学科重新进入。

当答案可以被无限生成，教育必须保护的不是无知，而是形成问题的时间；不是低效率，而是学习者在困难中取得自己的能力；不是教师的知识垄断，而是可追责、能修复也能放手的关系；不是过去的统治，而是有来源、有差异、可批判、可修复并可重建的文明记忆。记忆可以保存，继承必须由新主体参与。技术可以延展抵达，教育必须把判断、成长、关系和责任交还给真实的学习者。

锁定命题

教育不是把答案装进一个人，而是让一个人能够在未知中形成判断，并把自己的生命带入共同世界。AI 可以延展知识、练习与抵达；任何系统、学校或教师都无权把一个尚在成长的人写成已经完成的结论。

续读与共同研究入口

本文止于教育学、学习科学与代际承接能够独立承担的判断。续读不是把读者留在同一套文本中，而是当问题越过学习证据边界时，让真正相关的学科、竞争理论和现实验证重新进入。读者可以从自己的问题出发，不必按体系顺序阅读。

按问题继续，而不是按文件编号继续

方向	继续追问	建议续读／交接	为什么进入
上	为什么“学习效果”最终成为文明问题？	《AGI 时代的学科重新对账——学科边界、相互作用与共同研究议程》现行版（当前 V2.4）；文明跃迁五卷书。	看见教育结论与主体、权利、资源、制度、文化和公共权力共同运行后的整体后果。
下	这些判断怎样进入学校、公共教育与现实系统？	AI 教育制度设计；Human-COS、AGI-COS、公共窗口协议、国家级综合实验等相关现行文稿。	把支架撤除、儿童权利、责任链、退出和长期能力变成可验证的制度与工程条件。
左／横向	学习证据越界后，谁重新进入？	P03 认识论与科学哲学；P11 AI 与计算机科学；P05 伦理／法／政治哲学；P06 政治学；P09 社会／人类／文明研究。	分别处理证据强度、系统能力、权利资格、公共授权、关系文化与代际后果。
右／反向	如果不同意本文的人本边界，最强竞争路线是什么？	教师替代优先、AI-only 规模化、全量学习画像、数据驱动个性化、集中统一课程平台，以及“高绩效足以扩大教育权威”等竞争方案。	这些路线必须以长期学习、独立能力、公平、权利、关系和真实成本与本文正面对比，而不是被预先排除。
向前	哪些工作能够直接改变本文？	延迟保持与迁移实验；AI 撤除试验；真人支持／教师共驾 RCT；长期依赖与关系研究；学习者数据目的隔离；跨文化课程与代际承接研究。	用负结果、复现、强反例和真实制度后果迫使本文收缩、纠错、重构或失败。

如果主要关心学习效果，最值得继续做的不是再比较一次“有 AI/无 AI”的平均分，而是建立更完整的证据链：即时表现是否转成理解，理解能否迁移，迁移能否保持，撤除 AI 后能力是否仍在，面对冲突与未知时学习者能否形成自己的判断。任何一环出现反向结果，都应改变教育部署。

如果主要关心教师与 AI 的关系，研究也不应被压成“替代还是不替代”。可以分别检验 AI 作为知识工具、练习系统、智能导师、教师助手和评价辅助的净增量，同时把教师时间、专业能力、责任负担、关系质量和退出

能力纳入结果。若 AI 在某些功能上长期优于真人，应扩大相应功能；若依赖削弱学习者或公共教育能力，也必须允许系统退居。

如果主要关心儿童、文化与代际承接，需要继续进入儿童权利、社会学、人类学、文明研究与宗教／精神传统等相邻领域。保存得更多不等于继承得更好；翻译得更快不等于文化理解；被平台长期记住也不等于学习者应被历史数据定义。下一代必须拥有重新解释、拒绝、修复和创造的真实空间。

公开文库：<https://www.civitas.top/library/> 文明跃迁体系公众阅读与传播总入

口：<https://www.civitas.top/documents/p01-entry-006/> 五卷书入

口：<https://www.civitas.top/documents/p02-five-001/>

共同研究的入口不是“认同本文”。最有价值的参与，是提出更好的学习构念、更强的竞争解释、更长期的数据、更严格的撤除实验、更真实的文化与权利反例，或者证明某项边界没有必要。能够改变理论的专业工作，本身就是对文明跃迁研究最重要的参与。

不是为了让更多人读完文明跃迁，而是为了让更多人一起完成文明跃迁。

文明未来链 | 教育最终不是把人训练得更适应系统，而是把什么样的未来交给尚未完成的人？

愿景不是证据。下面的入口不用于证明本文已经成立，而是把本学科判断放进一个可以运行、生活与形成的未来文明中继续检验：读者既可以据此形成方向与向往，也可以拒绝现有图景，并用更强的专业事实、制度方案或现实结果改变它。

继续追问	建议进入	此时能看见什么	必须守住的边界
完整未来图景	《活的文明》三部曲统一出版合编版	让教育研究者看到学习、关系、制度与新文明形成最终共同指向怎样的生活世界。	未来图景不是教育效果证据，也不能成为统一课程目标。
孩子长大后进入怎样的生活	《活的文明·第二卷：未来文明的真实生活》	从成长、家庭、工作、关系与人生阶段判断教育究竟保护了什么能力与主体性。	一种理想生活不能被反向写成儿童的标准人生轨迹。
下一代怎样共同形成新文明	《活的文明·第三卷：新文明的形成》	继续观察教育、代际承接、社区与制度怎样让新一代拥有重新解释、拒绝和创造的空间。	代际传承不是价值复制，新一代拥有改变未来图景的权利。
学习怎样进入个人思想与行动	《思想跃迁》	从知识学习进入自我判断、生命经验与现实行动，检验教育是否帮助人成为能够独立思考的人。	思想转变不能被学校量化为人格分或服从指标。
AI 怎样成为人的能力外延	《星火全景高位说明》	观察 AI 怎样在授权边界内支持关心、学习、能力与真实关系，而不是把教师或学习者变成系统附属。	功能辅助不能自动取得教师资格、监护责任或评价权。

固定回流：[文明跃迁完整文库](#) | [公众阅读与传播总入口](#)

Human-COS 等尚未在统一公共映射中取得无歧义稳定单项页的入口，本稿只写明职责，不猜测 URL；需要时从完整文库进入当前公开版本。

附录 A 来源—主张—边界登记

表 A1 主要来源组的证据地位

来源组	本文使用的主张	等级	禁止外推
教育哲学[1-6]	教育涉及主动学习、经验、对话、新生、主体化与关怀。	A：文本事实；B：规范竞争。	任何一位思想家裁定所有教育或为原创命题背书。
学习科学[9-19]	提取、负荷、参与、动机、反馈和适度失败影响限定学习结果。	A：限定机制；B：跨情境解释。	短期表现代表理解、判断、人格或完整教育。
社会文化与关系 [7-8,26-29]	学习嵌入工具、共同实践、关系与文化传递。	A：多方法支持；B：机制争论。	人类教师本体排他性或所有学习必须面对面。
智能导师[21-25]	结构化系统可在特定任务和比较中改善结果。	A：旧 ITS 限定效果。	生成式 AI 普遍有效或完整教师替代。
生成式 AI 研究[30-35]	效果受护栏、任务、比较、测量和撤除条件影响。	B：新近、异质、快速变化。	单一实验或元分析成为跨场景定律。
国际规范 [36-45,49-51]	以人为中心、儿童权利、教育权、风险治理与公共知识。	A：正式文本；B：实施解释。	国际原则直接给出所有法域参数。
情绪与文化多样性[46-48]	情绪推断和生成内容同质化存在限定证据与技术风险。	A：情绪证据综述；B：生成研究。	由短任务证明人格或文明必然同质化。
未来 AGI 主体	体验、连续性与后果意识可能成为教育关系再审对象。	C：概念压力测试。	当前部署许可、教师资格、监护或政治权力。
文明跃迁哲学 ／规范承诺	意义多源、有限生命现实重量、能力不自动生成教育权威、未来主体证据开放。	F：公开哲学 ／规范承诺	写成学习科学、神经科学或 AI 工程已经证实的经验事实。

附录 B 十二项禁止性推导

- 从信息接入或正确解释，直接推出学习已经发生。
- 从提示内任务完成，直接推出撤除支架后的独立掌握。
- 从即时测验表现，直接推出长期保持、迁移或判断。
- 从点击、凝视、语气、情绪标签或满意度，直接推出人格成长。
- 从个性化准确，直接推出自主、尊重或有权锁定路径。
- 从教学功能，直接推出教育关系、教师资格或监护权。

- 从社会称呼与依恋，直接推出本体主体性或制度角色。
- 从教师资格或法律人格，反向推出道德完美、体验或政治资格。
- 从形成性反馈有效，直接推出资格认证、资源分配与社会选拔权。
- 从公开、可保存或技术可抓取，直接推出知识可训练、可生成和可商业化。
- 从危机期 AI-only 安排有效，直接推出稳定期关系降级正当。
- 从未来可能的新主体，追溯取消既有有限生命已经形成的教育关系与意义。

附录 C 外部审查清单

表 C1 投稿前学术与证据审查

检查项	通过标准	失败信号
陌生读者入口	不阅读体系文库也能复述问题、根基、术语和核心结论。	依赖内部编号、系统组件或未解释口号。
教育学对话	经典理论、学习科学、教师教育与儿童权利各在其范围内出现。	以通识概括替专业研究，或借权威背书。
AI 效果证据	任务、人群、比较、支架、测量、撤除和分配被说明。	以辅助产出、短期满意度或单一论文证明学习。
事实与规范	经验事实、竞争解释、体系判断与制度推导明确分开。	B/C 类写成稳定科学，或 F 类哲学／规范承诺被伪装为经验事实。
权力分层	功能、关系、资格、评价、监护、主体和权力分别举证。	能力、承认、权利、权力或主权自动互推。
可修订性	核心命题有反证、复核、阶段边界和撤回路径。	当前谨慎永久化，或未来开放成为提前授予。
哲学根基	意义多源、有限现实化、当前证据定位和未来开放完整保留。	死亡被写成唯一意义来源，或未来主体取消既有生命意义。
文本治理	无内部工程词、无伪精确参数、所有研究与规范可核验。	编号迁移代替重写，或制度建议假装全球通用。

生成式人工智能使用声明：本稿在资料与文献线索整理、结构设计、部分文字起草与润色、跨稿一致性检索、引用线索核查和版式检查中使用了生成式人工智能工具。生成式人工智能输出未被作为经验数据、权威来源或独立证据。研究问题、理论命题、规范判断、材料取舍、引文采用和最终定稿均由作者审定；作者对事实准确性、引文真实性、论证及全部结论承担责任。人工智能不列为作者，也不承担学术责任。正式投稿时，将按目标期刊要求补充工具名称、版本、使用日期和具体范围。

参考文献

- [1] 《论语·为政》2.15; 《论语·述而》7.8. 中国哲学书电子化计划. <https://ctext.org/analects/wei-zheng/zh> ;
<https://ctext.org/analects/shu-er/zh>
- [2] Dewey J. *Democracy and Education: An Introduction to the Philosophy of Education*. New York: Macmillan; 1916. Project Gutenberg edition. <https://www.gutenberg.org/ebooks/852>
- [3] Freire P. *Pedagogy of the Oppressed*. Ramos MB, trans. New York: Herder and Herder; 1970. 50th Anniversary edition, Bloomsbury Academic; 2018. <https://www.bloomsbury.com/us/pedagogy-of-the-oppressed-9781501314155/>
- [4] Arendt H. *The Crisis in Education*. In: *Between Past and Future*. New York: Viking Press; 1961. Current Penguin Classics edition. <https://www.penguinrandomhouse.com/books/299802/between-past-and-future-by-hannah-arendt-introduction-by-jerome-kohn/>
- [5] Biesta GJJ. *Good Education in an Age of Measurement: Ethics, Politics, Democracy*. Boulder, CO: Paradigm Publishers; 2010. <https://www.routledge.com/Good-Education-in-an-Age-of-Measurement-Ethics-Politics-Democracy/Biesta/p/book/9781594517914>
- [6] Noddings N. *Caring: A Relational Approach to Ethics and Moral Education*. 2nd ed. Berkeley: University of California Press; 2013. Original edition 1984.
- [7] Vygotsky LS. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press; 1978. <https://doi.org/10.2307/j.ctvjf9vz4>
- [8] Lave J, Wenger E. *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press; 1991. <https://doi.org/10.1017/CBO9780511815355>
- [9] National Academies of Sciences, Engineering, and Medicine. *How People Learn II: Learners, Contexts, and Cultures*. Washington, DC: National Academies Press; 2018. <https://doi.org/10.17226/24783>
- [10] Marton F, Säljö R. On qualitative differences in learning: I—Outcome and process. *British Journal of Educational Psychology*. 1976;46(1):4–11. <https://doi.org/10.1111/j.2044-8279.1976.tb02980.x>
- [11] Sweller J. Cognitive load during problem solving: Effects on learning. *Cognitive Science*. 1988;12(2):257–285. https://doi.org/10.1207/s15516709cog1202_4
- [12] Roediger HL III, Karpicke JD. Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*. 2006;17(3):249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- [13] Chi MTH, Wylie R. The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*. 2014;49(4):219–243. <https://doi.org/10.1080/00461520.2014.965823>
- [14] Dunlosky J, Rawson KA, Marsh EJ, Nathan MJ, Willingham DT. Improving students' learning with effective learning techniques. *Psychological Science in the Public Interest*. 2013;14(1):4–58. <https://doi.org/10.1177/1529100612453266>
- [15] Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*. 2000;55(1):68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- [16] Kapur M. Productive failure. *Cognition and Instruction*. 2008;26(3):379–424. <https://doi.org/10.1080/07370000802212669>
- [17] Sinha T, Kapur M. When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research*. 2021;91(5):761–798. <https://doi.org/10.3102/00346543211019105>
- [18] Black P, Wiliam D. Assessment and classroom learning. *Assessment in Education*. 1998;5(1):7–74. <https://doi.org/10.1080/0969595980050102>
- [19] Messick S. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*. 1995;50(9):741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- [20] Shulman LS. Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*. 1987;57(1):1–22. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- [21] Bloom BS. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*. 1984;13(6):4–16. <https://doi.org/10.3102/0013189X013006004>
- [22] VanLehn K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*. 2011;46(4):197–221. <https://doi.org/10.1080/00461520.2011.611369>

- [23] Ma W, Adesope OO, Nesbit JC, Liu Q. Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*. 2014;106(4):901–918. <https://doi.org/10.1037/a0037123>
- [24] Kulik JA, Fletcher JD. Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*. 2016;86(1):42–78. <https://doi.org/10.3102/0034654315581420>
- [25] Steenbergen-Hu S, Cooper H. A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students’ mathematical learning. *Journal of Educational Psychology*. 2013;105(4):970–987. <https://doi.org/10.1037/a0032447>
- [26] Roorda DL, Koomen HMY, Spilt JL, Oort FJ. The influence of affective teacher–student relationships on students’ school engagement and achievement. *Review of Educational Research*. 2011;81(4):493–529. <https://doi.org/10.3102/0034654311421793>
- [27] Tomasello M, Kruger AC, Ratner HH. Cultural learning. *Behavioral and Brain Sciences*. 1993;16(3):495–552. <https://doi.org/10.1017/S0140525X0003123X>
- [28] Caldwell CA, Millen AE. Experimental models for testing hypotheses about cumulative cultural evolution. *Evolution and Human Behavior*. 2008;29(3):165–171. <https://doi.org/10.1016/j.evolhumbehav.2007.12.001>
- [29] Mesoudi A, Whiten A, Dunbar RIM. A bias for social information in human cultural transmission. *British Journal of Psychology*. 2006;97(3):405–423. <https://doi.org/10.1348/000712605X85871>
- [30] Bastani H, Bastani O, Sungu A, Ge H, Kabakcı Ö, Mariman R. Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*. 2025;122(26):e2422633122. <https://doi.org/10.1073/pnas.2422633122>
- [31] Kestin G, Miller K, Klaes A, Milbourne T, Ponti G. AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*. 2025;15:17458. <https://doi.org/10.1038/s41598-025-97652-6>
- [32] Wang RE, Ribeiro AT, Robinson CD, Loeb S, Demszky D. Tutor CoPilot: A human–AI approach for scaling real-time expertise. *EdWorkingPaper 24-1056*. Revised November 2025. <https://doi.org/10.26300/81nh-8262>
- [33] Ma N, Zhong Z. A meta-analysis of the impact of generative artificial intelligence on learning outcomes. *Journal of Computer Assisted Learning*. Published online 1 September 2025. <https://doi.org/10.1111/jcal.70117>
- [34] Liu X, Guo B, He W, Hu X. Effects of generative artificial intelligence on K–12 and higher education students’ learning outcomes: A meta-analysis. *Journal of Educational Computing Research*. 2025;63(5):1249–1291. <https://doi.org/10.1177/07356331251329185>
- [35] OECD. *OECD Digital Education Outlook 2026: Exploring Effective Uses of Generative AI in Education*. Paris: OECD Publishing; 2026. <https://doi.org/10.1787/062a7394-en>
- [36] UNESCO. *Guidance for Generative AI in Education and Research*. Paris: UNESCO; 2023. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- [37] UNESCO. *AI Competency Framework for Teachers*. Paris: UNESCO; 2024. <https://unesdoc.unesco.org/ark:/48223/pf0000391104>
- [38] UNESCO. *AI Competency Framework for Students*. Paris: UNESCO; 2024. <https://unesdoc.unesco.org/ark:/48223/pf0000391105>
- [39] UNESCO. *Global Education Monitoring Report 2023: Technology in Education—A Tool on Whose Terms?* Paris: UNESCO; 2023. <https://gem-report-2023.unesco.org/>
- [40] International Commission on the Futures of Education. *Reimagining Our Futures Together: A New Social Contract for Education*. Paris: UNESCO; 2021. <https://unesdoc.unesco.org/ark:/48223/pf0000379707>
- [41] United Nations. *Convention on the Rights of the Child*. Adopted 1989; entered into force 1990. <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child>
- [42] UN Committee on the Rights of the Child. General Comment No. 25 (2021) on children’s rights in relation to the digital environment. *CRC/C/GC/25*. <https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>
- [43] United Nations. *International Covenant on Economic, Social and Cultural Rights, Article 13*. Adopted 1966; entered into force 1976. <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights>
- [44] United Nations. *Convention on the Rights of Persons with Disabilities, Article 24*. Adopted 2006; entered into force 2008. <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-persons-disabilities>

- [45] European Union. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Official Journal of the European Union. 2024. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj?locale=en>
- [46] Barrett LF, Adolphs R, Marsella S, Martinez AM, Pollak SD. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*. 2019;20(1):1–68. <https://doi.org/10.1177/1529100619832930>
- [47] Doshi AR, Hauser OP. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*. 2024;10(28):eadn5290. <https://doi.org/10.1126/sciadv.adn5290>
- [48] Shumailov I, Shumaylov Z, Zhao Y, Gal Y, Papernot N, Anderson R. AI models collapse when trained on recursively generated data. *Nature*. 2024;631:755–759. <https://doi.org/10.1038/s41586-024-07566-y>
- [49] UNESCO. Recommendation on Open Educational Resources (OER). Adopted 25 November 2019. <https://www.unesco.org/en/open-educational-resources/mandate>
- [50] CAST. Universal Design for Learning Guidelines version 3.0. Released July 2024. <https://udlguidelines.cast.org/>
- [51] UNESCO. Recommendation on the Ethics of Artificial Intelligence. Adopted 23 November 2021. <https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence>
- [52] Robinson CD, Gormley D, Ribeiro AT, Loeb S. Access is Not Enough: Human Support Improves Engagement with AI Tutoring. *EdWorkingPaper 26-1451*. Annenberg Institute at Brown University; June 2026. <https://doi.org/10.26300/pz7p-p388>