

# 人工智能时代还要考“《孔雀东南飞》主人公是谁”吗

## 人工智能是否颠覆了传统教育行业

李侠

近日,浙江某高校教师负责的3门课的卷面不及格率在30%以上。面对质疑,他在有关回应文章中列举了一道曾出过的试题,以此评价试卷的难度:“《孔雀东南飞》的主人公分别是谁和谁?”他说这道题的错误率在50%以上。

该事件引发了社会的广泛热议。其热议的核心在于在人工智能时代,对于那些可以快速检索到的知识,我们是否还需要像传统模式那样让学生背诵和记忆?换言之,现在的考试应该考什么?

### 理解才是力量

即便在人工智能赋能各行业已成潮流的当下,教育从本质、结构到机制也并未发生根本性改变,人工智能只是作为学习的辅助工具起作用,完全不能替代个体的主动学习。为此,我们不妨把教育的结构拆解开,看看人工智能到底能在哪个环节起作用,以及能起多大作用。

教育的结构通常分为三部分,即筛选、培养与输出。在这一链条上,教育的主要功能是通过教育的规划作用使学生形成特定的认知模式、习性/行为模式并完成初步世界观框架的搭建,同时在固定时间内对既定的培养目标进行测评,这个环节就是种类繁多的各种考试。

至此,我们大体了解了教育的完整流程。那么,在这几个环节中,人工智能是在哪个环节发挥作用的呢?

显然,人工智能的最大功用只是方便了知识的获取,还可以延伸到培养的末端,即知识测评。在其他环节,人工智能的作用有限。从这个意义上说,至少在当下,人工智能还没有颠覆传统教育范式。

回到那位老师的问题,这样的考题今天还有意义吗?在笔者看来,当下我们还是要辛苦记下《孔雀东南飞》的主角是谁,虽然人工智能/网络里可以轻易找到这些知识,但这并不意味着你的头脑中就有。

更何况,学习的过程是一个塑造个体认知图式的过程,如果没有点滴知识的缓慢学习,个体认知图式的形成过程将失去根基,因

为我们拥有的诸多人生经验都是通过亲力亲为获得的。由此,各种判断以及同理心才会被塑造并激活,就如同你即便看过无数爱情故事,也无法与真实的爱情相比一样。如果知识不能内化并嵌入到我们的认知图式中,我们如何作出各种价值判断呢?

澳大利亚哲学家杰克逊曾提出一个著名的思想实验“玛丽屋”。该实验说明,即便玛丽知道了无数关于红色的光学知识与生理学知识,对于红色的理解也无法与看到一个真实的西红柿相比,这在哲学上被称作“解释鸿沟”,其实质就是她缺乏对于红色的感受性。从这个意义上说,从“知道”到“理解”存在遥远的距离。

多年前,笔者曾把英国哲学家培根的名言“知识就是力量”改为“理解才是力量”,这一点在人工智能时代越发重要。在我们获得知识的工具与渠道愈加便捷与多样化的当下,也并未见有多少大师与重大成果横空出世。我们曾熟悉的传统学习模式被世人认为落伍了,其实不然,这种认知偏差反而导致我们离真正理解知识越来越遥远,甚至对那些常识知识不屑一顾。

更有甚者,有人则认为传统教育会扼杀人的创造性。事实上,科学史上的无数事例都会驳倒这些偏见,因为引发科学革命的大师们都是传统教育的产物,反而是新式教育的当下很少出现大师。

### 教育的数字化转型仍有边界

在此时代背景下,我们会对人工智能系统产生依赖并上瘾,久而久之,形成群体的认知路径依赖。这种依赖不可避免地导致群体的认知退化,因为我们很容易接受人工智能提供的答案,而不愿意亲自思考与探索。

照此趋势,人类行为所积累的数据信息会出现质量快速下滑的趋势。由于得不到高质量数据的大量训练,在低质量数据环境中运行的人工智能将出现功能退化,这种退化对依赖者产生反噬作用,从而在人类与人工智能之间形成一种双重无限退化的循环。

这个现象很好理解。众所周知,以大语言模型为代表的生成式人工智能本着“大力出奇

迹”的原则,围绕数据、算法与算力这三个维度开启技术迭代的进化之旅。如果数据质量持续走低,即便有优秀的算法及强大的算力加持,也不会使人工智能有明显的进步。

2023年3月,美国语言学家乔姆斯基在《纽约时报》上发文指出,人工智能和人类在思考方式、学习语言与生成解释方面的能力,以及道德思考方面有着极大差异。他提醒说,如果ChatGPT式机器学习程序继续主导人工智能领域,人类的科学水平以及道德标准都可能因此降低。

我们并不是要否认人工智能对整个社会产生的潜在深远影响,只是想指出,目前人工智能的作用被严重夸大了。这对于整个社会的有序运行与健康发展是不利的。

人类已经历过三次工业革命,每次都给人类社会的面貌带来了巨大改变,但那些改变仅限于物理层面。按照德国技术社会学家阿诺德·盖伦的说法,技术变革经历三个阶段——首先是工具阶段,其次是机器阶段,最后是自动机阶段,在这一阶段,主体的智力投入成为不必要的了。而我们目前正经历的人工智能革命,则要使人的思考能力也成为不必要的。这才是人们对这次技术革命喜悦参半的关键原因所在。

近日,有专家表示,传统概念中的大学可能将不复存在,越来越多的人会选择通过网络大学获得知识。在笔者看来,这种技术乐观主义有些过于自负了。未来其量只有少数较成熟的理工科课程可以采用此种模式,而大多数人文社科的课程都不适合,即便强行推进,效果也不一定乐观。

试想,一门课程如果由全世界水平最高的人在网上传授,那么人类思考的多样性与差异性也就会随之消失,这对于人类的未来到底是福还是祸?

笔者不否认人工智能会给当下的教育行业带来很多变化,但如果教育的基本功能结构没有发生改变,那么教育的数字化转型就是有边界的。毕竟人工智能只是手段而非目的,教育的一项重要功能就是培养学生的健全人格与世界观,而这一点人工智能是做不到的。

(作者系上海交通大学科学史与科学文化研究院院长)

编者按

近日,浙江某高校教师在网络上曝光,由于其教授的3门课程学生不及格率高于30%,学校要求该老师作出说明并提供整改方案。回复中,该教师列举了一道他出过的填空题以评价试卷难度:“《孔雀东南飞》的主人公分别是谁和谁?”据称,该题错误率在50%以上。

无独有偶,近日,谷歌前首席执行官(CEO)埃里克·施密特在美国斯坦福大学课堂上所进行的一场访谈中也被学生问及,在人工智能几乎可以取代程序员的当下,还有必要学习C++语言吗?

这使得我们不得不思考,人工智能时代,大学老师应教学生什么,又该考学生什么?如何定位课程和考试的意义与价值?考试的理念应做怎样的调整,形式又应如何变革?

不知道《孔雀东南飞》的主人公,在人工智能时代并不稀奇。人工智能颠覆了知识的存在方式和生成方式,学习、教学、考试也会相应发生改变。

一个人是否学习,掌握了规定的知识,是上个时代大学课堂考核的核心。一个人是否有自我驱动的学习意愿以及能力,将是人工智能时代的人成长的关键。

能让一个人获得幸福感的事,是挑战你感兴趣的事物。当下学生普遍最缺乏、也是人工智能带来的最重要改变之一是形成自我认知,包括个人兴趣、个人追求。大学应该针对这一变化做出积极改变。

### 关键在于自主意识觉醒

对于人工智能出现后的教学和考试怎么改革,目前有两种思路——赋能型思路和重塑型思路。现行教育过程核心有三步,第一步是教师教,第二步是学生学,第三步是考试。目前多数教改集中在赋能型思路,即在教育教学过程不变的前提下,怎样用人工智能让这三步变得更高效率。

然而,人工智能对教育的长远影响是重塑。因此,在人工智能时代,这三步本身可能没有意义,而是需要新三步来做学习支持,但我们在这方面的探索少之又少。

比如,一些高校教学大纲中规定的知识点,教师哪怕有一个点没讲到,都属于教学事故。但人工智能的到来会打破传授知识点的传统。在新体系中,学生学什么是由兴趣驱动的,以问题为导向,他会学习很多与某个问题相关的内容,且不由课堂提前设计,每人学到的内容也不相同。

人工智能时代,培养学生的能力是必须的。未来,几百人、几千人来一家公司做相同工作的情形会越来越少,更多的是每个人做不同的事情。独特性、个性化的前提是自主性。因此,自我意识、自主性这种高阶素养很重要。

但这并不意味着人工智能出现后,一个人不需要学习。当一个人的大脑没有足够知识储备时,可能连像样的问题都提不出来。因此,知识性考核不可能完全取消,而是取决于在什么场景中考核它,目的是什么。只不过知识的学习不需要完全在大学课堂进行。你可以在小憩时躺着学习,也可以在坐公交、地铁时静静地学习。

重点在于,课堂上学生不是学习了什么,而是有没有产生学习动机。高校更大的顾虑是学生未必会随时随地学习,后者可能更愿意打游戏、刷短视频。因此,问题的关键在于学生本身。他们经过十几年的正规学校教育后已被驯化,完全遵从老师安排的被动式学习导致其自主意识被极大弱化。

### 在真实场景中找问题

我们现在迫切需要一种教育理念和实践——这门课是开放的,不是教师布置给学生才去学,不是学完几个知识点才叫结课,而是要建立问题导向,即带着感兴趣的真问题学习这门课程,而非教师给学生“造”一个问题。这样做会最大程度激发学生的兴趣,一旦引起共鸣,他就会自觉地进入大量学习。

ChatGPT的出现对高校考核的冲击很大。我们发现,有些学生交上来的作业一看就不像自己所写,因为写得太好了。

比如关于苏州园林的古建筑保护,如果只是让学生谈一谈古建筑保护的思路,学生使用人工智能很容易获得答案。但如果让学生走进园林,经过实地调查提出一个自己感兴趣的题目,并和相关人员合作提供一套保护方案(例如拙政园的窗户保护),人工智能就只能泛泛而谈,此时考查的便是学生的真实水平。

题目来源于具象的真实场景中,有具体应用和创新性思考,这是对考核的一种调整。

在西交利物浦大学,我们给大一学生开设了一门5学分的课程,这门课没有教师在教室里上课,学生6-8人分为一个小组“真题真做”。比如一个与环境可持续发展相关的项目,学生不可能立刻找到答案,而是要把这个问题重新定义,做社会调查,将其细化后提出一个真实的问题。不管是社会调查、查文献,还是问教师、学慕

# 未来的考试可能是一篇反思性报告

张晓军

课,我们让学生自主开发解决方案,并鼓励项目落地。

然而,成败并不是关键。我们关注的是过程——学生彼此合作走向社会,关心真实问题,并把复杂问题简化为可以解决的问题,在问题驱动下,自主学习解决问题的知识、理论,并尝试落地的效果。在这个过程中,学生的训练和收获远超过教师给学生上20个小时课。

对于这种“做中学”的理念,有人质疑“还不如让学生直接去工作,不用上大学”。事实上,学习与实践还是有区别的。学习允许失败,而且应以过程中学生的不断反思作为考核重点,以人的成长为核心。这与真实世界的实践不大相同,因为几乎没有人会允许你做5个项目而“不论成败”。

### 反思性评价是未来走向

高阶能力、素养能否通过考试来考查?笼统地说很难。但通过丰富的实践培养并考核高阶能力,我认为才是未来的方向。人工智能可以生成丰富的知识,但它对于人类未来的价值不在于人能学会多少,而在于人能否在实践中利用人工智能生成内容,并在此过程中实现自己对社会的独特价值。

如今,所有学生坐在教室里拿到相同的卷子,这是脱离具体场景的抽象考核,很难唤起学生的反思与学习的自主性。

大学四年应该是学生从对自我长期发展不是特别清晰到比较清晰的阶段,以及学习的自主性从不太强到比较强的阶段。因此,衡量的标准也要发生变化。

反思性评价在一定程度上已经存在。比如,大学里都在做课程思政,如果发一张卷子让学生谈爱国主义精神、社会责任感是什么,即便回答得很完美,也难以考查出学生爱国主义精神、社会责任感的高下。而通过写反思报告,让他接触到的某个社会现象,如何理解这一问题、他为解决该问题做了什么、做后的感受等,便体现了他的情感共鸣,即社会责任感的提升。这类的考查通过反思报告更加合适。

通过人工智能大数据建模,可在一定程度上反映学生的自主性。比如,看学生所学知识是教师指定的内容,还是他自主学习了很多与课程相关的内容。如果是前者,说明学生自主性不够强,后者则与自主性紧密关联。

看学生的学习时间,如果集中在期中、期末考试时段,主要是为了通过考试,说明他的自主性不够强;如果学习时间分散在平时,不在意考试的因素,在一定程度上可以体现他的自主性。

无论是反思性学习还是自主性学习,都是基于前一段学习的总结和回顾,并对下一阶段的提出有价值的启发,而不是非得给学生一个分数。这一点同样值得注意。

(作者系西交利物浦大学未来教育学院执行院长,本报记者温才妃采访整理)

# 人工智能时代,如何提高大学课程及格率

文双春

近日,国内某高校教师因其所负责的3门课程的卷面不及格率都在30%以上,被学校要求作解释说明并提供整改方案。该教师在网络上公开发表说明文章,引发关注和热议。

众所周知,不及格率高的课程在全世界高校司空见惯。笔者不禁联想到与此对应的另一更普遍现象——课程分数通胀,即学生普遍得到了比应得分更高的分数。

分数通胀的课程是否意味着比不及格率高的课程的教学质量更高?似乎很少有师生这么认为。但由于分数通胀带来“皆大欢喜”,所以往往既可免去教师诸如作解释说明、提供整改方案等麻烦,更有可能使其在学生评教系统中收获同样的分数。

不可否认,在不降低考试难度、不放松评分标准(放水)的前提下,几乎所有高校和教师或多或少面临课程不及格率高的问题。因此,如何在保障质量的前提下提高课程及格率,是值得所有高校和教师探讨的课题。

### 单靠课堂教学无法提高课程及格率

大学常常从课堂教学中寻找课程及格率低的原因:一是教师授课能力和水平有问题,二是学生不进课堂或不专心听讲。但笔者认为没有抓住根本和要害。

大学的课堂教学基本沿袭教师讲、学生听的讲课模式,但这种模式不断受到质疑和抨击,其中重要原因是教育者往往以自己接受教育的方式教育受教育者。

任何教学都必须转化为学生的学习成效,否则,充其量是一场表演秀。早在1946年,著名学习专家爱德加·戴尔就提出“学习金字塔”理论,并指出听课是所有学习方式中效果最差的。

2019年,诺贝尔物理学奖得主卡尔·韦曼基于其团队的研究,尖锐地指出:几乎可以肯定,几个世纪以来的讲课一直是无效的。

该团队的一项研究表明,如果纯粹是听课,那么一门课程由经验丰富的资深教授讲,还是由没有任何教学经验的研究生讲,学生的学习效果没什么区别。

如果讲课一直是无效的,学生进不进课堂、用不用心听讲还重要吗?

德国卡塞尔大学的工作人员研究了到课率与学习成绩之间的关系,发现总体而言,学生的到课率和成绩之间存在正相关,但多为弱相关;随着学习技术的变化,大学的到课率似

乎失去了它的重要性,因为学生不必上课也可以获得课程材料。

### 学生想学习并积极参与时,才可能提高及格率

一所大学的同专业、同课堂基本由高考成绩相差无几的学生构成,意味着这些学生的前期基础、智力水平、学习能力、考试技巧等没有什么差别,但为何一些学生成绩优异,另一些学生却不及格?

实践表明,学习动力是最主要的因素。认知科学也告诉我们:只有当学生想学习并有强烈的参与欲望时,学习才会发生。而学生是否想学习和想学到什么程度,又与学生对“意义或目标”的认识紧密相关。

直白一点说,学生往往用现实或功利主义眼光来衡量——想保研或考研的学生,跟保研或考研相关的每一门课程成绩都优异或至少不会不及格;想考公考证的学生,完全靠自学也能学好相关课程。

当今,大学中找不到“意义或目标”,从而缺乏学习动力的学生不少,他们往往构成课程不及格率的“基本盘”。

美国哈佛大学发布了一份关于年轻人(18~25岁)心理健康的调查报告。报告指出,36%的年轻人报告有焦虑情绪,29%的年轻人报告患有抑郁症。年轻人情感挣扎的原因有多种,其中高达58%的年轻人报告他们的生活缺乏“意义或目标”。

一种教育或一门课程教学是否有成效,最终看其在多大程度上改变或塑造了学生。教育有如打铁,你首先得把铁烧红,然后才能下锤打出你期待的效果,否则就是打冷铁,你很辛苦,铁很痛苦,而铁没有任何改变。

大学如果不能帮助学生找到“意义或目标”,老师如果只是传授知识而不能让学生理解知识蕴含的真正意义,那么大学和老师在“打冷铁”。

### 人工智能时代,手机是最有成效的个性化课堂

孔子曰:“不愤不启,不悱不发。”《西游记》中菩提祖师说:“不遇人传妙法,空言口困舌头干。”

古代圣贤告诉我们,教学需要把握时机,否则怎么教都没用。现代心理学的心流理论证

明了这种思想。

心流是人们在做什么事情时表现出的全神贯注、投入忘我的状态。个体在达到心流状态时才是最有效的学习。心流体验源于任务挑战性与个体技能水平的平衡;当任务的难度与个体的技能水平相当时,才有可能产生心流体验;任务过易会导致无聊,过难则会产生焦虑。

另外,根据美国著名教育学家戴维·索恩伯格所著《学习场景的革命》,教学实践取决于教学环境,善用学习场景,才能更有效地学习。

戴维认为,人类一直在“营火”“水源”“洞穴”和“生活”4种独立学习场景中学习,它们分别代表从故事讲述者或者专家那里学习(讲属于这种场景)、同伴之间互相学习、学习者通过反思和内省的方式进行独立学习,以及将理论付诸实践的应用型学习。

戴维指出,要让学生获得真正、持久且有意义的知识,应减少学生对“营火”场景的依赖,让其在4种场景中都有足够的学习时间,并可在4种场景中自由切换。

学习效果与时机和场景的密切相关性决定了教学为什么很难,因为教师不可能知道每个学生学习发生的确切时机、确切环境,从而因时、因地、因材施教。这也是“讲课一直是无效的”重要原因。

笔者认为,借助网络和人工智能,可以突破这种困境。关键要做到两点:一是把教学主导权交给学生,教师从教育者转变为学生的学习的辅助者;二是支持学生把手机变成最有效的个性化课堂,培养学生成为“因材施教”的主动学习者。

今天,许多学生不愿进课堂,即使进课堂,也很少抬头看PPT大屏,而是低头刷手机。在笔者看来,任课教师对此不仅不要失落,反而要支持,当然前提是学生在刷手机——如果学生有更好的途径掌握我希望他们学到的东西,我有什么理由不支持呢?

特别是人工智能正在超指数级地增强自身能力,从传统的机器学习不断向自我进化、富有洞察力和自主解决问题演化,已经并将更深入地改变大学的教学。不难想象,人工智能将比任何人类教师都更智能地使学生达到心流状态,为学生提供或创造最适配的学习场景。

笔者相信,当每一个学生都善用手机学习时,大学课程及格率将不再是问题。(作者系湖南大学物理与微电子科学学院)