

深度学习的知识加工机制与教学条件

黄 梅 黄希庭

摘要：深度学习是对知识进行精制编码和深度加工的学习，表层学习是深度学习的基础，在一定教学条件下，表层学习可以向深度学习转化。通过对基于原型表征的信息流模型、集中注意力嵌入过程模型、自适应共振状态机制模型、理性思维的自适应控制模型等认知模型进行研究认为，激活学生长时记忆阈值水平的情境学习、基于共激活点的大概念系统教学、基于知识本质理解的科学思维方法学习、陈述性知识程序化的科学探究活动、目标导向的元认知反思等教学条件，能够促进深度学习。

关键词：深度学习；知识加工；教学条件

在教学中，我们常常面临以下问题：为什么学生总是记不住或者记住了却不会用？如何纠正学生的错误概念？如何辨别学生是否听懂了？学生是否进行了有效学习？怎样引导学生进行创新性思考？除了布置作业，还有其他方法让学生巩固新知识吗？等等。这些问题都与深度学习（deep learning）有关。在知识的学习过程中，如果学生经常使用重复练习、重写笔记、死记硬背等表层记忆学习的方法，而不去深入理解、思考知识的本质，他们往往不能获得高质量的学习效果。

“深度学习”概念源于人工神经网络研究。本茨（Bentz）最早直接使用“深度学习”一词，且明确将“深度学习”定义为一种能力，这种能力包括智慧、情感、技能等，并具有一定的生长性，能够进行有效迁移。^[1]深度学习第一次应用于教学领域是在1976年。瑞典教育心理学家马顿（Marton）和塞尔乔（Saljo）联名发表的

《学习的本质区别：结果和过程》一文中阐述了深度学习和表层学习（surface learning）这两个概念。他们在对学生学习过程的实验研究中发现，学生在阅读同一篇文章时，存在两种取向的学习方式：一种取向是基于对文章的理解，概括出文章大意，并结合自身进行思考；另一种取向专注于可能会被提问的部分，并试图去复述这些内容。于是，他们认为，建立在理解、联系和自主建构基础上的高水平认知学习方式为深度学习，以机械记忆为主的为表层学习。亚历山大（Alexander）认为，当一个人处于不同水平的专业知识阶段时，知识加工的深度会有所改变。深加工最适合处于中高水平专业阶段的人，但对专业能力低的学习者却并非如此。某些策略（如表层策略）并不会在专业知识中消失，但是会减少，并继续与深加工策略协同工作。^[2]美国教育传播与技术协会在2004年对深度学习重新下了定义，指出深度学习需要通过对知识进行理解和

基金项目：西南大学宜宾研究院“宜宾市双城协议保障科研经费”科技项目资助项目“基于深度学习的基础教育提升体系”（XNDX2022020017）；西藏自治区教育科学研究重点课题“西藏中学化学教师应用国家优质教学资源的能力提升对策研究”（XZEDIP240016）。

迁移,最终引发高阶思维和认知;深度学习的目的在于促进学习者形成批判性思维,提高学生的创新意识。^[3]综上所述,我们认为,深度学习是对知识进行精制编码和深度加工的学习。表层学习是深度学习的基础,在一定教学条件下,表层学习可以向深度学习转化。

一、深度学习的知识加工机制

随着认知心理学、脑神经科学的发展,认知模型的研究经历了从基于原型表征的信息流模型、集中注意力嵌入过程模型、自适应共振状态机制模型到理性思维的自适应控制模型等不断完善的过程,对知识加工机制的研究在不断深入,这些研究逐步破解了人类学习与记忆的奥秘,促进了教育科学的发展。特别是对深度学习的研究,解决了人工智能中自然语言理解、知觉处理、一般学习等重要问题,促进了人工智能的进步。

(一) 基于原型表征的信息流模型

阿特杰森(Atkinson)和希福林(Shiffrin)提出,外界信息进入记忆系统后,经历了感觉记忆、工作记忆和长时记忆三个阶段。^[4]外部信息先输入感觉记忆,但多数信息很快就会消失,部分信息会以听觉的、口语的或书面语言的形式重新编码,被选择性地加工成可理解的事物或特征进入工作记忆。感觉记忆的信息转化为工作记忆的过程受到注意力的控制。注意力越集中,就越能够把信息加工成工作记忆。但工作记忆的信息也会很快消失,信息必须通过编码、检索、复习等深加工过程才能完成从工作记忆到长时记忆的转化。通常情况下,长时记忆被分为语义记忆(事实记忆)、过程性记忆(动作技能和日常行为)和情景记忆。语义记忆贮存了关于对象的事实和其相互关系的知识,语义知识像一个概念图一样执行,概念之间通过节点相互关联并表征彼此之间的关系;在过程性记忆系统中,过程性知识显现了一个“如果—那么”的规则,知识在这里被深度加工,建立起认知的因果关系模型;情景记忆则贮存过去经验中具体的例子,如果遇到相似情景,这些记忆就能被再次使用或用于开发新的情景性或过程性知识。

长时记忆对注意力的要求相对较低,不仅能

被动地发挥作用,而且能主动地和策略性地发挥作用。^[5]在教学中,教师可以有策略地通过情境线索激活学生先验知识、集中学生注意力及引导知识线索,以减轻注意力的负担,抵消注意力减少而引起的衰变和干扰。同时,也可以通过组织大概念学习、陈述性知识程序化、建构知识因果关系模型等知识深加工策略,尝试在长时记忆中创建牢固的结构或项目组,当有更多的知识进入长时记忆时,就可以利用长时记忆的激活来抵消对注意力的需求。

(二) 集中注意力嵌入过程模型

科恩(Cowan)认为,所有感觉输入都与长时记忆相关,大脑的活动基本上就是一束束神经元的激活,需要进行精致加工才能进入长时记忆,或者需要有认知冲突的变化才导致不自觉地注意。集中的注意力部分被自发地控制,部分因为环境的变化而不自发地改变。^[6]科恩的嵌入过程模型假设贮存与加工都要使用注意,但某些挑战性的加工必须检查注意的容量,否则将被放入临时记忆中。

工作记忆是一种对信息进行暂时性加工和贮存的记忆系统,它为推理、深入理解等高级认知任务提供了加工时所必需的信息和临时性存储空间。工作记忆的容量是有限的,当新的信息到达时,最老的和最不相关的项目将被删除,以避免信息过载。当项目很长时间没有使用时,也将被删除。^[7]这个理论告诉我们,神经元的激活需要能量,这些能量来源于外界的刺激。教学刺激的强度与知识加工方式决定着是否能有效地激活学生原有的知识经验,而学生原有的知识经验彼此之间是否进行了有效的联系,是否因经常使用而巩固,决定着学习的效率。

(三) 自适应共振状态机制模型

自适应共振状态机制模型将认知模型划分为内容、策略和元认知,最后将这三部分内容整合在一起,构建了自我调节学习和认知负荷相互作用层级模型。在这个模型中,感觉记忆和工作记忆嵌入了长时记忆之中,内容层级、学习策略层级和元认知层级分别主导着学习任务需要和学习表现、学习策略的支持与应用、元认知线索和元认知调节的作用。

此模型认为,除了遗忘衰减的信息,被选

择的信息在感觉记忆、工作记忆和长时记忆之间的相互作用有两条路径。第一条路径是激活的感觉记忆与长时记忆的内容发生了自适应共振状态,感觉记忆的信息在共振状态中被放大,在无认知负荷、无意识、本能的、自动化的过程中与长时记忆中的内容相互作用,检索和提取长时记忆的内容进行新信息的学习,新的信息进入长时记忆对原来的知识结构图式进行补充与完善。第二条路径是当感觉记忆的信息与长时记忆的信息不匹配或者激活过强时,感觉记忆的信息进入工作记忆,工作记忆的信息必须通过有目的、有意识和有认知负荷的精致加工才能进入长时记忆,与长时记忆中的信息发生作用。^[8]在这个过程中,策略模块根据反馈和是否实施成功等条件选择表层策略还是深度策略,使用陈述性、程序性还是条件性策略知识。元认知模块在此过程中充当了标准的设置、计划、监督和控制者的作用,不断地调节、选择、匹配整个系统中人、任务和策略之间的关系。^[9]所以,高效的教学活动是知识的共振激活、深加工、有目标的元认知反思三者有效结合的过程。合适的教学情境对先验知识的激活共振效应会让教学变得轻松而高效,知识通过积累、重复、练习、强化、修正、分析、比较、推理、组织、模型化等深加工会促进知识的长时记忆。

(四) 理性思维的自适应控制模型

脑科学的研究为理性思维的自适应控制 (adaptive control of thought-rational, ACT-R) 模型提供了理论基础。ACT-R 已经进化到 ACT-R7 版本,发展为包含多个模块的理论,并且为解释这些模块如何整合起来以产生连贯的认知提供了理论基础。

ACT-R 模型是基于大脑生理机制的模型。在 ACT-R 模型中,视觉模块用于识别视野中的客体,手动模块用于控制手,陈述性模块用于从记忆中提取信息,支持学习、存储和对知识或与情境特定事件有关记忆的检索(情景记忆)以及语义记忆。目标模块用于保持对当前目标和意图的追踪。中央产生式系统 (production system) 是 ACT-R 体系结构的重要程序性系统,用于实现这些模块行为的协调。子符号过程系统主要作

用是为长时记忆中存储的陈述性知识块赋予一个数值,称为陈述性知识的活跃度。知识块的活跃度越大,越容易被检索和激活。

基底神经节和相关的连接被认为在 ACT-R 中实施产生式规则,对应大脑组织中的前额叶皮层—纹状体—丘脑环,执行模式识别、冲突解决、控制动作的功能,即通过具有模式识别功能的纹状体、具有冲突解决功能的苍白球和控制动作执行与产生的丘脑进行程序性记忆的学习。支持程序性记忆系统的主要神经结构是一组皮层下结构,称为基底神经节、小脑等。与陈述性记忆不同的是,程序性记忆更适合学习顺序上或概率上的结构化信息。通常程序性记忆系统学习很缓慢,为了学习一项知识或技能,需要多次复习或重复地接触这些信息。然而,一旦获得了信息,新的知识和技能可能会在没有意识的情况下使用。一个经常被引用的例子是我们在学会知识或技能后,在没有意识的情况下自动执行新的运动技能。重复运动技能(如伸手去拿东西)的行为将在没有意识的情况下迅速执行。^[10] ACT-R 模型告诉我们,有目标的学习最高效,知识越用越灵活,手脑并用才是学习的本质。所以,教师在教学中首先要让学生明确学习目标,注重理性思维和动作技能的结合,让学生通过动脑、动眼、动嘴、动手,充分调动大脑的感官系统模块,通过多感官的协同学习完成陈述性知识向程序性知识的转化,促进程序性知识的自动化。

二、深度学习的教学条件

(一) 激活学生长时记忆阈值水平的情境学习

科恩认为,一些信息被充分地激活或使用正确的方式激活,才能成为个体有限容量注意的部分。一个大脑组块的激活值与基线水平激活值(反映其过去知识的总体可用性)、连接激活值 (associative activation, 反映其和当前背景的相关性) 以及引起注意的强度有关。^[11] 当学生解决问题时,他们首先注意到与先验知识的不一致性或导致自己失败的信息,这些信息使他们开始注意区分先前的知识,通过努力参与到对信息的更深入的加工当中。虽然他们有可能最终无法解决问题,但学习过程中面临的困难或错误会迫使学生进行一种极限的认知努力,会不断修正自己原

来的认知结构,并在此基础上进行新的学习。所以,我们认为,只要进行了深入的认知努力,同时激活了先验知识,就意味着学习的发生。

在知识的学习中,教师应设计能够激活学生长时记忆阈值水平的教学情境。这种教学条件有如下特点:(1)通过情境线索为新旧知识建立有效的逻辑联系,激活先验知识;(2)利用教学情境的新异性、不确定性和预测性,引发学生形成认知冲突;(3)教学情境关联知识的意义和价值,激发学生的学习动机;(4)注重目标引导和不同年龄阶段学生认知负荷的特点,避免专长逆转效应。以上这些教学条件,能够促进学生长时记忆的激活水平达到阈值,从而激发感觉记忆与长时记忆产生自动化的共振状态,达到高效教学的目的。

(二) 基于共激活点的大概念系统教学

根据激活扩散模型,当某一概念受到刺激时,该节点即被激活,并沿着各条连线向四周扩散,到达下一个节点,从而激活其余关联概念。^[12]对人工智能层级神经网络的研究表明,共激活图中的中心节点与深度神经网络中各自的神经元之间存在关联。如果从模型中移除具有高网页级别中心性的神经元,深度神经网络的准确度将大幅降低,而如果移除具有低网页级别中心性的神经元,准确度将保持稳定。这表明,共激活图中的中心节点可能代表深度神经网络中的重要神经元。^[13]这对教学的启示是,如果我们知道这个核心的共激活概念节点,教学的时候就可以利用其激活关联着的其他重要知识的节点,来影响整个知识网络的加工深度和连续性,达到举一反三的目的。

基于共激活点的大概念系统教学是促进对知识形成更加连贯、系统理解的教学,这种教学条件有如下特点:(1)让学生形成基于共激活点核心知识的知识体系,建立具体概念之间的因果联系;(2)将知识结构与学生的认知层次和认识逻辑结合,达到概念的高度统摄性;(3)打破学科间的壁垒,形成学科内、学科间、学科与生活之间的高度集成的跨学科知识联系。

(三) 基于知识本质理解的科学思维方法学习

基于知识本质理解的科学思维方法学习能够促进学生独立思考能力和创造能力的培养。大量研究证明,只有在深度信息处理的背景下,大脑

才会活动。所以,教学要注重对知识本质的高认知水平探究活动,让学生学会用证据进行解释、分析、比较、归纳、演绎,用实验进行验证与推理,深入探讨知识的表面特征与内在本质之间的联系,挖掘现象背后的科学意义和规律,不断更新、评价与修正原有的认知模型,促进概念在新条件或情境下应用和迁移。

基于知识本质理解的科学思维方法学习的教学条件有如下特点:(1)让学生像科学家一样,学会对所产生的问题进行假设、比较、分类、设计实验方案、实施实验、获取证据、解释或建构模型,对数据和证据用严密的逻辑进行准确推理,最终达到问题解决的目的。通过这种科学思维方法的训练,让学生不仅关注“是什么”的问题,更关注“为什么”“如何做”的问题;(2)让学生会从宏观、微观、符号、曲线、公式、模型等不同表征认识知识的本质,形成透过现象看本质的理性思维方式;(3)重视教学过程中逻辑的、理性的认知加工过程。训练学生基于事实证据或实验结果,运用形式逻辑规则对知识本身及其因果关系进行定性推理和定量推断,用合理的理由支持所阐述的观点,基于证据说明数据与结论之间的联系,能够运用控制变量法研究设计科学的实验。

(四) 陈述性知识程序化的科学探究活动

心理学研究认为,程序性知识获得的心理机制是产生式,是从陈述性知识转化为程序性知识的结果。^[14]最初贮存进长时记忆中的知识结构是不完善的,需要在不断提取的刺激过程中完善。长时记忆中贮存的信息,在外界的刺激下提取出相应的知识块,在这样重复发生的过程中,知识被反复提取应用,使条件项与行动项的联结增强,知识块的活跃度增强,命题网络逐渐转化为产生式系统。这个过程即条件项逐渐概括化和分化,完成陈述性知识向产生式系统的转化。^[15]

陈述性知识程序化的科学探究活动教学条件有如下特点:(1)陈述性知识和程序性知识学习后需要立即练习应用并不断重复巩固,一旦信息巩固并自动化,新的知识和技能就可能在没有意识的情况下自动执行;(2)将动作技能与智力技能相结合进行学习,用实验和计算同时训练来促进陈述性知识的程序化;(3)运用顺口溜、因果

推断等联系定位策略促进陈述性知识向产生式系统转化；(4) 陈述性知识可以程序化，程序性知识反过来可以推断陈述性知识；(5) 将论证与建模作为展开科学探究的重要部分，培养学生熟练运用简单联系、复杂联系、抽象公式、贝叶斯概率化四层不同复杂水平的知识关系模型来解释、理解、分析和运用科学知识，让长时记忆中的认知图式更加优化，促进知识的应用和迁移。

(五) 目标导向的元认知反思

元认知层次越高，推理和批判性思维水平也越高，知识建构就越好。库恩认为，有效的元认知策略思维的挑战是协调理解任务和理解潜在的策略。^[16]尼克森（Nickerson）认为，从长时记忆中检索信息是通过自动/随机检索和更具战略性/定向搜索之间的平衡来实现的。检索的过程就是在长时记忆中查找和选择信息的循环过程，其中检索到的信息又作为进一步检索的线索。整个检索过程中既有定向的部分，也有随机的部分。定向部分是指由个人直接控制的策略性过程，包括建立检索计划、选择和生成适当的线索来检索记忆，以及各种监控策略和决定继续检索与否。随机部分指检索过程的概率性。定向控制过程对于从长期记忆中成功检索信息至关重要，有目标的检索会增加成功的概率。^[17]

目标导向的元认知反思学习教学条件有如下特点。(1) 注重目标导向反馈和反馈的阶段性和时效性，要给学生多元的反馈机会并及时反馈。比如，教师讲过一个知识点之后，学生马上进行几个小练习，教师及时对学生存在的问题进行纠正反馈后再进入下一个知识点的学习。(2) 让学生运用标准范例获取反馈信息，促进自我的学习，养成主动接收反馈的素养。比如，给学生一个标准答案，要求学生根据标准答案来进行自评，学生可根据评价指标了解自己知识的掌握情况，在检查、对比标准答案的过程中，思考更加独特的解题方法。(3) 让学生在解决问题后及时总结，因为过去的成功经验会让学生未来选择它的频率增高。比如，要求学生反思“我今天做对了哪些？我今天没做对哪些？我今天要改善哪些？”“你还有什么其他新方法可以解决这个问题吗？”(4) 给予学生在社会性关系中进行元认知反馈的机会，让学生在知识的交流交换过程中获

得一定的启发。比如，学生彼此交换作业，相互批改作业，让学生之间互相评价。(5) 开发挑战性自我反思、纠正和调节的机会。比如，教师给学生自评的答案，但是这个答案不一定正确，让学生在评估教师答案和自己答案的时候作出正确的判断，这种方式是在同伴互评的基础上把原认知水平再往上提高了一步。

参考文献：

- [1] BENTZ V M. Deep learning groups: combining emotional and intellectual learning [J]. *Clinical sociology review*, 1992 (1): 71-89.
- [2] DINSMORE D L. Toward a dynamic, multidimensional research framework for strategic processing [J]. *Educational psychology review*, 2017 (2): 235-268.
- [3] 刘志波, 李阿琴. AECT2004 定义解读 [J]. *电化教育研究*, 2004 (12): 44-48.
- [4] SPENCE K W, SPENCE J T. *The psychology of learning and motivation* [M]. New York: Academic Press, 1968: 89-195.
- [5] RHODES S, COWAN N. Attention in working memory: attention is needed but it yearns to be free [J]. *Annals of the New York academy of sciences*, 2018 (1): 52-63.
- [6] 黄梅. 知识激活阶段的教学条件 [J]. *中小学教材教学*, 2021 (4): 54-58.
- [7] GZIGLER I, WINKLER I. Unconscious memory representations in perception: processes and mechanisms in the brain [M]. Amsterdam: John Benjamins Publishing Company, 2010: 1-93.
- [8] WIRTH J, STEBNER F, TRYPKE M, et al. An interactive layers model of self-regulated learning and cognitive load [J]. *Educational psychology review*, 2020 (4): 1127-1149.
- [9] UNSWORTH N. Examining the dynamics of strategic search from long-term memory [J]. *Journal of memory and language*, 2017 (1): 135-153.
- [10] ANDERSON J R, BOTHELL D, LEBIERE C, et al. An integrated theory of list memory [J]. *Journal of memory and language*, 1998 (4): 341-380.
- [11] ANDERSON J R, BOTHELL D, BYRNE M D, et al. An integrated theory of the mind [J]. *Psychological review*, 2004 (4): 1036-1060.

(下转第 55 页)