

信息技术赋能科学素养评测：发展脉络、研究热点与未来趋势

Information technology-enabled scientific literacy assessment: developmental history, research hotspots and future trends

吴国静¹, 王靖²

江南大学江苏“互联网+教育”研究基地

2994237780@qq.com

【摘要】 面对科学教育中科学素养评测科学化、高效化、个性化的需求，传统评测方式的短板愈加明显。因此，信息技术赋能的科学素养评测发展历程、现状和未来走向成了科学教育领域的关键议题。为此，对 1950-2025 年国内外科学素养评测实证研究成果进行系统性文献综述，讨论了科学素养评测方式的转变、热点和未来趋势。研究发现：科学素养评测方式经历了纸笔测试、计算机静态评测、基于计算机的人机交互三个阶段。研究焦点集中于基于多场景表现性评价和融合多模态数据的科学素养评测。最后，本文从积极融入生成式人工智能、关注本土创新等五个方面做出了展望。

【关键词】 科学素养评测；科学素养；信息技术；评测方式；科学教育

Abstract: In the face of the demand for scientific, efficient and personalized scientific literacy assessment in science education, the shortcomings of traditional assessment methods have become more and more obvious. Therefore, the development history, current status and future direction of information technology-enabled scientific literacy assessment have become key issues in the field of science education. To this end, a systematic literature review of the empirical research results on scientific literacy assessment at home and abroad from 1950 to 2025 was conducted to discuss the transformation, hotspots and future trends of scientific literacy assessment methods. It is found that scientific literacy assessment methods have gone through three stages: paper-and-pencil test, computerized static assessment, and computer-based human-computer interaction. The research focuses on scientific literacy assessment based on multi-scenario expressive evaluation and integration of multi-modal data. Finally, the paper makes an outlook in five aspects, including active integration of generative artificial intelligence and attention to local innovation.

Keywords: Science literacy assessment, scientific literacy, information technology, assessment methods, science education

1. 引言

当前，建设科技强国，发展新质生产力已成为国家战略的核心命题，科学教育不仅承载着知识传递的功能，更肩负着培养科学素养、激发创造潜能、培育创新人才的重要使命。2022 年我国教育部最新修订的《义务教育科学课程标准（2022 年版）》为科学教育的开展提供了指导方针，并提出了将科学素养作为学生发展的重要目标。国务院 2021 年颁布的《全民科学素质行动规划纲要（2021-2035 年）》提出“科学素质是国民素质的重要组成部分，是社会文明进步的基础”（国务院，2021）。对科学素养进行评测是检验科学教育效果不可或缺的环节。目前已有的评测主要是以书面形式为主，仅以文字和图片呈现给学生（陈

睿智等，2024），以分数体现学生的科学素质和能力情况。然而随着技术的发展，人工智能、大数据等为实现科学素养评测的动态化、个性化提供了契机。

在此背景下，本文提出以下研究问题：在已有的科学素养评测研究中，技术赋能的科学素养评测的经历了哪些发展阶段？当前新技术不断涌现，科学素养评测研究焦点集中在哪些方面？新兴技术与教育评测深度融合的背景下，科学素养评测将呈现怎样的发展趋势？为此，本文对科学教育中以技术支持的科学素养评测的发展进行分析与解读，通过分析科学素养评测的各阶段特点，总结以技术支持的科学素养评测研究的热点和未来发展趋势，以进一步把握科学素养评测的发展轨迹以及未来研究取向，为后续的科学素养评测的研究和实践提供有益参考。

2.科学素养概述

二战后，军事上形成了以美国和苏联为首两大阵营对峙的局面，科学研究团体认识到科学技术对于国防的重要价值，而科学技术的提升离不开公民对技术的理解和支持。美国国家科学基金会（National Science Foundation, NSF）在1950年明确表示“科学进步需要获得公众对科学教育和科学研究的支持”（AT Waterman, 1960）。

1952年，美国化学家、教育学家、哈佛大学前校长J.B.柯南特第一个提出科学素养定义。柯南特指出，科学素养是一种明智地选择专家的能力，以及能够“聪明地与那些正在发展科学并应用科学的人交流的能力”（Richard, 1998）。尽管没有对其意义与内涵进行深入研究，但是这从普通教育的层面对科学素养进行了定位，为后续的科学素养相关研究指明了方向。1958年赫德在基础教育领域采用了科学素养这一概念，并将其解释为“理解科学且能把其应用在社会中”。可见，由于时代背景的需要，20世纪五六十年代提倡的科学素养是科技精英人才、职业科学家所具备的科学知识、理解科学事业的能力。

60年代中期，美国的佩勒（Pella）等人从六个方面对科学素养这一概念进行了综合性的概括（魏冰，2000）。佩勒等人的研究成果建立起了科学素养的基本框架，为科学素养的深入研究奠定了基础。

20世纪70年代，经济全球萧条和科技进步带来的环境问题，国家重新关注个体的需求和发展，科技政策也将重点从国防转向社会服务。在此背景下，科学素养的内涵也进一步得到丰富。具备科学素养的人能够“理解科学、技术和社会经济发展其他方面的关系”（NST Association, 1971）。到了80年代，科学素养不在局限于理论层面的讨论，而是作为教育目标明确提出“能够理解科学、技术、社会之间的相互影响并能在日常生活中运用这些知识做出决策”（TNST Association, 1982），这也标志着科学—技术—社会教育正式拉开序幕。这一时期，“科学”和“技术”被置于同等地位，提升公民的科学素养不再是为了培养科学家，而是面向全体学生，实现“科学为大众”。

科学素养评测以明确科学素养的概念为基础，并在此基础上围绕科学素养的结构框架展开。1983年Miller提出了科学素养三要素模型“（1）科学术语与概念的词汇量；（2）对科学过程的理解；（3）科学技术对个人和社会影响的了解”（Aikenhead G S & Ryan A G, 1992）。这也是后来米勒公民科学素养三维理论模型的雏形。

20世纪90年代末，人类迈入了信息时代，互联网、云计算、大数据等信息技术的蓬勃发展，短视频、公众号等新媒介不断融入公众的生活，打破了书籍、报纸和电视作为传播科学信息的唯一途径的局面。科技信息的获取、传播速率和内容产出在数字媒介技术迭代的驱动下便捷性不断提高，尽管公众接触低信度、伪科学的概率和风险增加了，但是媒体平台构建

的点赞、评论和转发的异步交互场景，已经深度重构和影响了公众对科学的态度和理解。在这高度信息化的社会背景下，科学素养内涵也不断拓宽。美国国家科学院联合多个组织机构编写的《科学素养：概念、情境与影响》强调了“面对信息洪流，科学素养需要的是整合并解释信息的能力，以及进行反思和评价的能力（凯瑟琳·E·斯诺，肯妮·A·迪布纳，2020）。大规模国际科学素养评测项目也在随着时代的变化完善评测内容，经济合作与发展组织（OECD）发布的 PISA2025 框架明确将“评估科学数据的可信度与缺陷”纳入核心能力维度。可见，随着科技的飞速发展，新兴事物不断涌现，日新月异的变革周期不断缩短，个体只有在掌握科学知识和技术的基础上独立批判地思考和发现问题、辩证地分析问题，才能做出合理决策来解决问题。

3.研究设计

3.1 研究方法

本研究采用系统性文献综述法（Preferred Reporting Items for Systematic Reviews and Meta-Analyses，简称 PRISMA），对国内外科学素养评测相关文献进行分析，以总结技术支持的科学素养评测发展脉络，精准把握当前该领域的研究热点，并基于现有研究对该领域未来发展进行展望。在明确研究问题的前提下，通过检索、筛选、合格和纳入 4 阶段对技术支持的科学素养评测的相关文献进行筛选。

3.2 文献获取

本文从 Web of Science 核心数据库中，使用数据库的高级检索功能以“Science Literacy Assessment or Scientific Literacy Assessment”为主题词进行检索，文献类型为“ARTICLE”，检索截止时间为 2025 年 3 月，共检索到科学素养评测相关文献 982 篇。

选取中国知网（CNKI）数据库的作为国内科学素养评测研究的数据库来源。以“科学素养评测”为主题，检索到论文共 59 篇。中外文数据库累计获取文献量 1041 篇。

3.3 文献筛选

为了保证文章的可靠性和质量，本研究制定了文献筛选标准，根据标准逐一对文献进行筛选。纳入与排除标准如下表所示。

表 1 文献筛选标准

编号	纳入标准	排除标准
C1	文献收录期刊是否是 SSCI、SCI、CSSCI 级别	非上述期刊
C2	研究主题是否为“科学素养评测”	研究主题不相关
C3	研究内容含有完整的研究设计、科学素养评测的工具或评测方法	无完整研究设计，不涉及科学素养评测工具或评测方法
C4	研究类型是否为实证研究	研究类型非实证

献的标题、关键词和摘要根据制定的文献筛选标准对文章进行初步筛选。其次，通过阅读全文对文献进行二次筛选。最后，利用“滚雪球”的方法检索纳入更多符合标准的文献，最终选出符合标准的文献共计 41 篇，文献筛选过程如图所示。

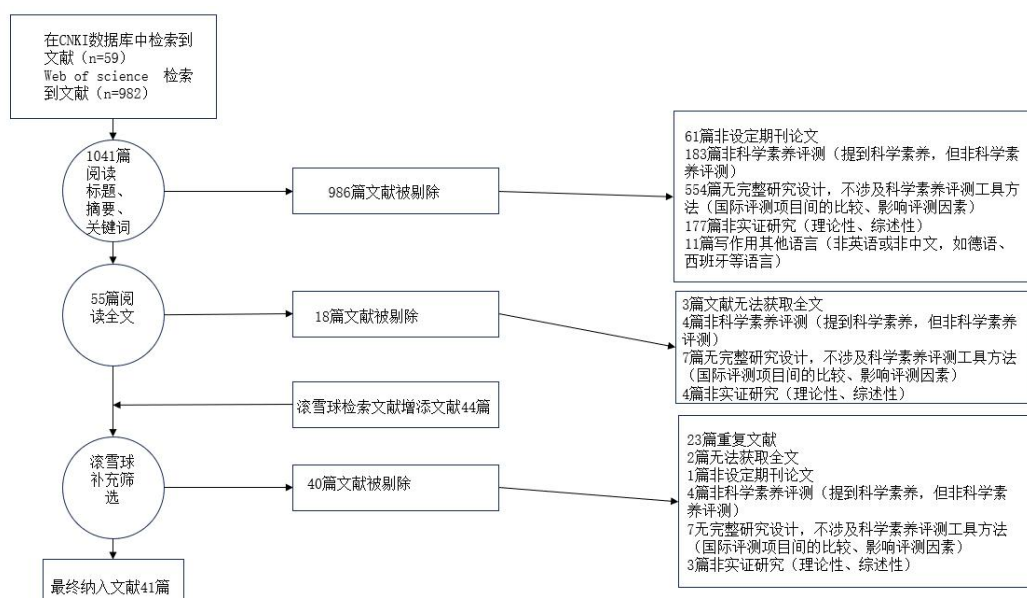


图 1 PRISMA 流程图

4. 研究结果和讨论

4.1 技术支持的科学素养评测的历史演进

最早对科学素养进行评测可以追溯到 1957 年美国开展的公众科学素养测评项目“公众对科学的理解”。这不久后教育评测领域就引入了计算机技术，但这之后的 20 年间，基于计算机的测试主要在课堂环境下用于学生的陈述性知识的回答判断（Shute, Valerie J. & Seyedahmad Rahimi, 2017）。虽然引入计算机技术是教育评测领域上的一大创新性突破，但是该阶段并没有充分利用和发挥计算机的优势。试题类型只有选择题和填空题，计算机的反馈局限于作答结果的正误判断，被试与计算机之间的交互限制于简单的鼠标点击。

80 年代末 90 年代初，国际上陆续开展科学素养评测项目，并利用评测结果推动科学教育改革和科技政策制定。这期间，科学素养评测主要是通过静态纸笔测验的形式进行。尽管纸笔测验便于实施，但纸笔测验存在着一系列问题，不仅难以评测学生的高阶思维，还脱离真实情境，不能评估出学生在真实情况下观察的能力（Shavelson, R. J., Baxter, G. P., & Pine, J, 1994）。

90 年代末起，研究人员 Baker 和 Mayer 等人开始使用计算机来评估更复杂的认知技能。科学教育领域也逐渐突破纸笔评测局限，引入了利用人与计算机交互开展评测的思想。美国教育测量服务中心的梅斯雷弗（Mislevy, R. J.）等研究者于 1997-1999 提出了“证据中心设计”（ECD）理论，该理论指导了美国国家教育进展评估（National Assessment of Educational Progress, NAEP）、国际数学和科学成就趋势研究（Trends in International Mathematics and Science Study, TIMSS）和国际学生能力评估项目（Program for International Student Assessment, PISA）等科学素养大规模评测项目中的基于信息和通信技术（Information and Communication Technology, ICT）任务的设计。

自此，人机交互理念走进大规模科学素养评测项目。2006 年，PISA 项目首次尝试利用人机交互的形式开展科学素养评测（OECD, 2010）。随后在 2009 年，NAEP 使用了操作性任务和计算机交互任务两种新题型评估学生的科学素养（NAEP, 2023）。这个时候尽管人机交互试题投入科学素养评测的时间较短，但其自动化和高效性还是受到了众多研究者的密切关注。

Quellmalz 等人依据提炼的基于计算机模拟的科学评估原则，围绕初中生态系统开发了 12 道题，通过 1566 名学生完成评估，得出了互动评估比静态评估在区分学生的探究能力方面更有优势的结论(Quellmalz, E. S. , Davenport, J. L. , Timms, M. J. , Deboer, G. E. , Jordan, K. A. , & Huang, C. W. , et al,2013)。

21 世纪以后，人机交互开展科学素养评测的研究不断深入。吴心楷等人开发了基于多媒体的科学探究能力评估工具，情境化开发了 114 道科学人机交互试题，并通过 1066 名学生的参与检验该工具上人机交互试题的有效性(Kuo, C. Y. , Wu, H. K. , Jen, T. H. , & Hsu, Y. S. ,2015)。2015 年 PISA 首次正式在大规模的科学素养评测中使用人机交互试题，在随后的 PISA2018 和 PISA2022 继续沿用了该评测形式。2019 年 NEAP 全面引入人机交互的测评题。TIMSS 也从 2019 年将评测形式从纸笔过渡到利用计算机进行评测。

上述描绘的技术支持的科学素养评测历史演进关键节点如下图所示。

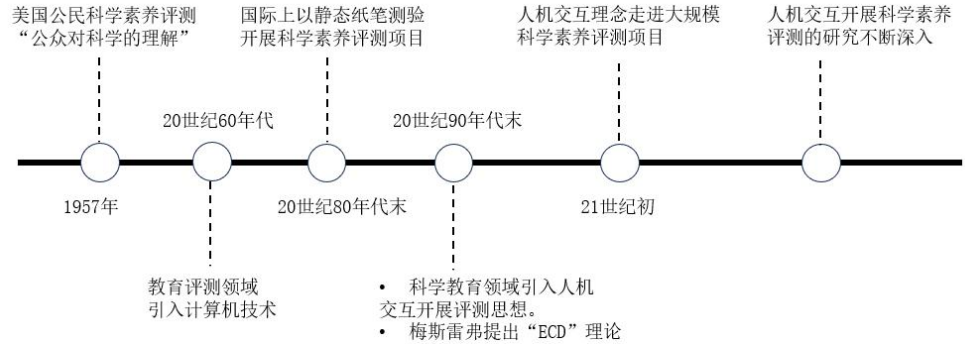


图 2 技术支持的科学素养评测历史关键节点

4.2 技术支持的科学素养评测研究热点分析

随着技术的发展，技术支持的科学素养评测正朝着个性化、智能化、更科学化评测的方向发展。在这一阶段，学者们不断探索和创新，尝试将多种技术和工具融入科学素养评测，期望提高科学素养评测的准确性和有效性。

“以证据为中心”的评测设计模式（evidence centered design, ECD）是系统性进行评价设计的模式，该模式已经成为教育评价领域主要的研究方向和应用领域之一。该模式明确指出教育评价本质上是基于学生的表现获取关于学生能力的证据进而对学生的知识、技能和成就进行推理（冯翠典，2012）。我国研究者在借鉴国际大规模评测项目和以证据为中心 ECD 模型理念的基础上，将科学素养评测方式从纸笔测验逐步过渡到计算机交互。陈慧珍等研究者借鉴 NAEP 计算机交互评测技术，围绕“温室效应与植物光合作用”设计并实施评测方案，施测结果显示，该评测技术能够充分考察学生的知识、能力素养等水平（陈慧珍等，2018）。刘枳衫、林婷婷同样设计并实施了基于计算机交互和动手实操任务的科学素养评测（刘枳衫，2016；林婷婷，2016）。除了基于计算机的交互，还有学者将传感器也融入科学素养评测中，以实现对学生的实操评测。庞谦竺在深入分析 PISA 和 NAEP 科学素养评估框架的基础上，设计了基于计算机和传感器的九年级科学素养评估方案（庞谦竺，2019）。周海涵同样将计算机平台、传感器技术融入到科学素养评测，设计了基于信息和通信技术（Information and communication technology, ICT）的“地球的运动”科学素养评测系统（周海涵，2022）。目前，PISA 已经采用自适应技术精准定位不同地区和群体的科学素养水平差异，动态匹配评测项目，满足个性化评测需求。也有研究者尝试将自适应技术的引入科学素养评测。通过在评测中根据学生的能力水平，动态调整评测项目的难度和内容，实现高效化、个性化科学素养评测(李小红等，2024)。

4.2.1 融合多模态数据的科学素养评测

在教育评测领域，充分借助数据的创新趋势来深化教育评价的趋势所在。随着人工智能技术、大数据、物联网等新兴一代技术的不断成熟，数据赋能于教育评价逐渐成为常态化。在此背景下，融入多模态的科学素养评测悄然兴起，并吸引了众多研究者进行尝试与应用。陈信宇通过对学生在科学素养评测中的概念任务和动手操作中引入眼动追踪技术，探究学生的注意力分配与认知差异，将客观的神经教育学数据融入科学素养评测（陈信宇，2020）。Andy Nguyen 等人借助生理信号、音频数据和视频数据用于探究光和声音物理概念协作学习中的社会调节过程，进而全面了解科学学习的过程，并利用多模态数据来调节、评估学习效果（Nguyen, A., Järvelä, S., Rosé, C., Järvenoja, H., & Malmberg, J., 2023）。

4.2.2 基于多场景表现性评价的科学素养评测

表现性评价能够反映学生习得知识并运用知识来解决问题，在这一过程中，重点评估学生的问题解决能力、高阶思维能力和实际应用能力，因此许多高学业成就的国家将其纳入课程和测试当中（张殷等，2024）。由上述的定义可知，表现性评价在评估批判性思维、创新能力等高阶能力方面具有巨大潜力。在科学素养评测上，表现性评价体现为进行实验设计、分析数据、得出结论，即完成一项实践探究，或者是解决实际问题，或者形成一个作品。具体来说有以下几个表现：设计情境化任务。学生在完成任务的过程中，需要读取相关资料，设计实验并解释现象。如 PISA2025 样题 1 “温室效应”中通过阅读文字和图表，寻找支持“二氧化碳排放量上升会导致气温上升”，并解释哪些信息不支持该论点（张殷等，2024）。模拟科学实验环境的交互式科学任务。这些任务通常在基于计算机的虚拟环境中进行一系列实验操作，实现真实问题的解决。NEAP、PISA 和 TIMSS 目前都已经采用了基于计算机的交互式测评。动手实验任务。评测过程中学生需要利用科学仪器和材料完成实验任务，与此同时，研究者依据一定的标准进行评估（张殷等，2024）。

4.2.3 数据驱动的智能化科学素养评测

随着技术的发展，人工智能、大数据等新技术的应用不断深入，新技术赋能科学素养评测使得精细多元的评价方式和伴随式数据采集逐渐成为现实。刘邦奇等认为，智能技术与教育评价相融合，驱动数据采集过程化、分析多模化、评价分析反馈及时化，教育评价的数字化、智能化、专业化得到显著提升（刘邦奇等，2021）。在科学素养评测领域，不少学者借助人工智能技术，尝试实现智能化评测。

Wilson 等利用机器学习实现学生科学论证以及绘制的模型的自动评分，研究结果显示人类批改的分数和机器给出的分数具有高度一致性（CD Wilson 等人，2023；zhai、He 等人，2022）。孔艺和林静借助机器学习的自动评分工具 c-rater-ML，利用图尔敏模型分析逻辑结构分析学生科学论证内容，准确率达 90%（孔艺，林静，2021）；王晶莹等人借助模糊神经网络技术对影响学生科学素养的相关因素进行挖掘，根据研究结果对科学素养的培养提出建议（王晶莹等，2024）。

聚焦当下信息技术的发展，随着人工智能在教育领域的应用，自动化测评及多模态人机之间的协同与交互持续受到了研究者的高度关注。这一现状既是智能技术驱动科学素养评测从经验驱动向数据驱动转型的结果，也是人机协同评测模式在实际教育场景中的应用需求。

4.3 对未来我国科学素养评测研究的展望

运用新技术改进和优化测评方式已成为当前教育质量监测评价改革的重要趋势（李小红等，2024）。21 世纪以来，技术赋能科学素养评测涌现出的一系列研究热点与问题，是我

国教育评测和科学教育发展方向抉择的重要参照与影响因素。研究者应当面向未来教育,紧跟时代发展的脚步重新审视科学素养评估。

第一,打破学科融合,多方参与,建设评测共同体。高等院校继续发挥其在科学素养学术研究方面的优势,通过不同研究机构、研究者之间的沟通合作,确保以技术支持的科学素养评测研究质与量均衡发展。其次,增强各学科之间的融合,打破学科壁垒,从教育学、科学、计算机、传播学等多学科视角进行研究,实现高水平的科学素养评测创新。

第二,确保科学素养评测基础夯实并紧密追踪热点,确保科学素养评测满足时代需求。首先,科学素养评测相关的基本问题,如科学素养的概念,在新的时代背景和理论视角下需要进行反思和重构,增添时代赋予的意蕴,紧扣人才培养需求。其次,从文献关键词的整体分析来看,科学素养与核心素养之间的关系,以及落实到教学一线,二者如何作用与影响也需要进一步探讨。

第三,在与国内外已有宝贵经验对话与借鉴的同时,关注自身本土化创新。既要吸收引进国际技术支持的科学素养评测相关研究成果,又能立足于本土特色与国情,使得评测方式能够迁移与落地。在科学素养评测工具方面,在研究、参与国际评测项目的同时要尝试自主设计、开发和应用技术支持的科学素养评测方案与工具。尽管目前由于设备环境等条件,科学素养评测仍然是基于纸笔测验进行开展,但由于纸笔存在的仅聚焦于知识层面、无法记录作答过程等一系列不足,基于计算机的人机交互和数字化、智能化评测工具成为科学素养评测的重要途径。在设计及使用上述方案与工具的过程中,既要吸收 PISA、TIMSS 相关经验,如整个评测流程的设计理念、评测框架的梳理与构建、试题的编制、评测结果的分析方式等等。同时也要关注课标中学生的学业质量要求并在设置问题情境时积极融入中华优秀传统文化。

第四、积极融入生成式人工智能,推动评测方式更加智能化。以 ChatGPT 为代表的人工智能生成式内容(artificial intelligence generated content, AIGC)在教育领域受到了高度关注。与传统数字技术不同,生成式人工智能技术,是指具有文本、图片、音频、视频等内容生成能力的模型及相关技术。生成式人工智能主要赋能于评测资源整合、试题命制、评测实施与评测结果的分析反馈(国家互联网信息办公室等,2023)。AIGC 作为一种新兴技术,其用于科学素养评测尚处于积极探索阶段,虽然产出了一定的研究成果,但在实际评测中并未具体实施。因此,研究者在将 AIGC 应用于科学素养评测时,既要看到其广阔的应用前景,也要关注到应用中面临的诸多挑战,应当加强实验试点,积极探索将生成式人工智能融入科学素养评测的机制和路径,推动科学素养智能化评测体系的建立与完善。

参考文献

- 陈睿智,谢晓雨,罗莹.(2024).科学教育测评中人机交互试题的发展、特征与启示.中国考试,(08),79-88. <https://doi.org/10.19360/j.cnki.11-3303/g4.2024.08.009>
- 陈信宇.(2020).眼动技术在科学素养评测中的应用研究(硕士学位论文,东南大学).硕士 <https://link.cnki.net/doi/10.27014/d.cnki.gdnau.2020.001842doi:10.27014/d.cnki.gdnau.2020.001842>
- 陈慧珍,刘枳杉,柏毅.(2018).高中生物学素养评测方案的设计与实施.教育测量与评价,(08),49-55. <https://doi.org/10.16518/j.cnki.emaec.2018.08.008>

- 冯翠典.(2012).“以证据为中心”的教育评价设计模式简介.上海教育科研,(08),12-16.
<https://doi.org/10.16194/j.cnki.31-1059/g4.2012.08.003>
- 国务院关于印发全民科学素质行动规划纲要(2021—2035年)的通知.中华人民共和国国务院公报,(19),12-20.2025-4-28 取自 https://www.gov.cn/gongbao/content/2021/content_5623051.htm
- 国家互联网信息办公室等 (2023).《生成式人工智能服务管理暂行办法》. 2025-4-28 取自 https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm
- 凯瑟琳·E·斯诺,肯妮·A·迪布纳(2020).科学素养:概念、情境与影响(裴新宁,正太年译).北京:中国科学技术出版社.
- 孔艺,林静.(2021).基于机器学习的科学论证自动评分个案研究.中国考试,(08),61-68.
<https://doi.org/10.19360/j.cnki.11-3303/g4.2021.08.009>
- 李小红,李珍,王克志.(2024).PISA的新进展分析及其启示.教育科学,40(03),31-37.
<https://doi.org/10.3969/j.issn.1002-8064.2024.03.005>
- 林娉婷.(2016).学生科学核心概念评测方案研究(硕士学位论文,东南大学).硕士
https://kns.cnki.net/kcms2/article/abstract?v=_dzes7vZToNws8h6lHVBnOF5GLcJTN8wFGQOQM6NUA_xrRVHFe1JhyUKIyVxXHwZIKrYcv1ODh9qfxwbGpuX9Y8d84uYQDqg3ruXGweb0REjBgjFGK8lAHVnb_lDvicqNtlBVN3uL2o03GIn_f69zucH520G-4qZ7KQ_Sgf6NCPSLx37k0Ifz4uizyyUxWIP9dsGOVMKWs=&uniplatform=NZKPT&language=CHS
- 刘邦奇,袁婷婷,纪玉超,刘碧莹 & 李岭.(2021).智能技术赋能教育评价:内涵、总体框架与实践路径.中国电化教育,(08),16-24. <https://doi.org/10.3969/j.issn.1006-9860.2021.08.003>
- 刘枳杉.(2016).学生科学素养评测方案研究(硕士学位论文,东南大学).硕士
https://kns.cnki.net/kcms2/article/abstract?v=_dzes7vZToP-F1vK6jAwCPouZVR87oL6QUXItKATeY7cZjU0GvvIUMZ-X_uN5v5JS9n7GJcofGBzd4ZsG69NEbvSgUcyRpQXZ9pNYCa0GQgNN12XSeA2kxb6-XRv4x4u5M450Yc_PFHdPTnqkHi6krOTfYmHcomX-uclieXWqdbwS5k_9lkWrHGFZZGxXTPfqymITSznWec=&uniplatform=NZKPT&language=CHS
- 庞谦竺.(2019).九年级科学素养评测方案设计(硕士学位论文,东南大学).硕士
<https://link.cnki.net/doi/10.27014/d.cnki.gdnau.2019.003165doi:10.27014/d.cnki.gdnau.2019.003165>
- 魏冰.(2000).“科学素养”探析.比较教育研究,(S1),105-108.
- 周海函.(2022).基于 ICT 的科学素养评测方案开发(硕士学位论文,东南大学).硕士
<https://link.cnki.net/doi/10.27014/d.cnki.gdnau.2022.001532doi:10.27014/d.cnki.gdnau.2022.001532>
- 张殷,吴柳燕,罗星凯.(2024).聚焦科学高阶思维的科学表现性任务设计研究.天津市教科院学报,36(06),3-14. <https://doi.org/10.3969/j.issn.1671-2277.2024.06.001>
- Association, N. S. T. . (2010). Principles of professionalism for science educators. national science teachers association position statement. National Science Teachers Association, 4.
<https://www.nsta.org/nstas-official-positions/principles-professionalism-science-educators>
- Association, T. N. S. T. . (1982). Science-technology-society : science education for the 1980s. The NSTA Position Statement. <https://scholarworks.uni.edu/istj/vol20/iss2/11>

AG Aikenhead . (2010). The development of a new instrument: 'views on science—technology—society' (vosts). *Science Education*, 76(5), 477-491. <https://doi.org/10.1002/sce.3730760503>

Kuo, C. Y. , Wu, H. K. , Jen, T. H. , & Hsu, Y. S. . (2015). Development and validation of a multimedia-based assessment of scientific inquiry abilities. *International Journal of Science Education*, 37(14), 2326-2357. <https://doi.org/10.1080/09500693.2015.1078521>

National Assessment of Educational Progress .(2012). *Science in Action: Hands-On and Interactive Computer Tasks From the 2009 Science Assessment*. https://doi.org/10.1007/978-94-007-2150-0_67

Nguyen, A., Järvelä, S., Rosé, C., Järvenoja, H., & Malmberg, J. (2023). Examining socially shared regulation and shared physiological arousal events with multimodal learning analytics. *British Journal of Educational Technology*, 54(1), 293 - 312. <https://doi.org/10.1111/bjet.13280>

Organization for Economic Co-operation and Development .(2010). *Pisa pisa computer - based assessment of student skills in science: (édition complète - isbn 9789264082038 - en angl. seulement)*. Source OCDE Science et technologies de l'information, volume 2010, i-132(132). <https://doi.org/10.1016/j.compedu.2020.103964>

Quellmalz, E. S. , Davenport, J. L. , Timms, M. J. , Deboer, G. E. , Jordan, K. A. , & Huang, C. W. , et al. (2013). Next-generation environments for assessing and promoting complex science learning. *Journal of Educational Psychology*, 105(4), 1100-1114. <https://doi.org/10.1037/a0032220>

Shute, V. J. , & Rahimi, S. . (2017). Review of computer - based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning*, 33(1), 1 - 19. <https://doi.org/10.1111/jcal.12172>

Shavelson, R. J. , Baxter, G. P. , & Pine, J. . (1991). Performance assessment in science. *Applied Measurement in Education*, 4(4), 347-362. [https://doi.org/10.1016/S0191-491X\(99\)00021-8](https://doi.org/10.1016/S0191-491X(99)00021-8)

Waterman, A. T. . (1960). National science foundation: a ten-year resume. *Science*, 131(3410), 1341-1354. <http://doi.org/10.1126/science.131.3410.1341>

Wilson, C. D. , Haudek, K. C. , Osborne, J. F. , Buck Bracey, Z. E. , Cheuk, T. , & Donovan, B. M. , et al. (2024). Using automated analysis to assess middle school students' competence with scientific argumentation. *Journal of Research in Science Teaching*, 61(1). <http://doi.org/10.1002/tea.21864>

Zhai, X., He, P., & Krajcik, J. (2022). Applying machine learning to automatically assess scientific models. *Journal of Research in Science Teaching*, 59(10), 1765 - 1794. <https://doi.org/10.1002/tea.21773>