

题库型游戏评测批判性思维能力研究

冷静^{1,2} 路晓旭²

(1. 华东师范大学 上海数字化教育装备工程技术研究中心, 上海 200062;
2. 华东师范大学 教育信息技术学系, 上海 200062)

[摘要] 作为21世纪学习者的必备技能,批判性思维能力的培养和测评变得越来越重要。本文拟突破传统的纸笔测试,设计并开发一款测量中小学生学习批判性思维能力的工具。近年来,游戏化测评技术成为大数据时代评价学习的重要方式。游戏化测评可将观察到的学习行为对测评对象实现即时反馈。本研究设计了题库型游戏测评工具,将游戏化元素与批判性思维能力指标相结合,让测评对象在游戏情境中作出相应决策,通过在线行为了解其能力水平。研究发现,测评对象的批判性思维能力水平处于中等水平;得分高的测评对象会花更多的时间研读游戏情境中出现的重要信息,如例题等。本研究使用虚拟游戏元素使测评对象投入到任务情境中,以更真实、更快速的方式实现对批判性思维能力的评估,同时也期望利用游戏化测评技术推动大数据背景下其他高阶思维能力测评的发展。

[关键词] 游戏化测评技术;游戏设计与开发;题库游戏

[中图分类号] G434 **[文献标识码]** A **[文章编号]** 1007-2179(2020)01-0082-08

一、问题提出

为了适应21世纪社会的需求,沟通、协作、创造力、创新、批判性思维和解决问题的能力已成为个人不可或缺的必备技能(Miri et al., 2007; Trilling & Fadel, 2009; Yemenici, 2016)。其中,批判性思维被认为是一种高级思维,是科学素养的集中表现(褚宏启, 2016)。批判性思维能够增强学习者辨别信息真伪的能力,且基于多元观点辩证地进行问题分析、思考和决策(冷静等, 2018)。很多研究指出,在教育教学中培养批判性思维能力非常重要(Bassham et al., 2010; Facione, 2013; 玛雅·比亚利克等, 2019)。加强批判性思维的培养,需要对学

习者的批判性思维进行测评,以提出培养批判性思维的对策。然而,目前针对批判性思维的测评工具形式单一,主要以纸质版问卷或量表为主,存在一定局限性,比如题目枯燥且题量多,不适于对中小学生学习进行测评。

近年来,游戏化学习研究广受关注。游戏化环境可以呈现真实的情境,提高学习者的参与性,弥补纸质测评的不足。特别是题库型游戏(Item bank game)可以将游戏化元素和批判性思维能力测量指标相结合,通过测评对象在游戏中的行为实现对其能力的评估。本研究设计并开发了一款测评游戏,测量中小学阶段(四年级以上)学生的批判性思维能力,主要考察以下四个方面的能力:归纳、推理、观

[收稿日期] 2019-05-31 **[修回日期]** 2019-12-20 **[DOI编码]** 10.13966/j.cnki.kfjyyj.2020.01.009

[基金项目] 上海市科委科研计划“工程中心2017-2020能力提升项目”(17DZ2281800);认知智能国家重点实验室开放课题“基于大数据的区域教育发展画像研究”(iED2019-Z02)。

[作者简介] 冷静,博士,华东师范大学教育学部教育技术学系晨晖学者,研究方向:计算机支持的协作学习、批判性思维能力培养、教育数据挖掘(jleng@deit.ecnu.edu.cn);路晓旭,华东师范大学教育学部教育技术学系硕士研究生,研究方向:学习分析、教育游戏。

察(判断可信度)以及辨别假设。

二、文献综述

(一) 批判性思维的定义及测评

经济合作与发展组织启动的“核心素养”项目,提出将批判性思维放在核心素养概念框架中(钟启泉,2016)。美国发布的《21世纪学习框架》明确提出批判性思维的重要性。欧盟发布的核心素养框架指出,就个人、社会能力与学习能力而言,学习者需要具备批判性思维能力。中国核心素养课题组(2016)提出的中国学生发展的核心素养也包含批判性思维的培养。未来社会需要具备21世纪能力的人才,批判性思维是其中的核心能力(Ab Kadir, 2017; Alnofaie, 2013; Bedir, 2016; Dwyer et al. 2014; Li, 2016; Stupple et al., 2017)。

研究者未能对批判性思维的定义形成统一意见。美国哲学家约翰·杜威(John Dewey)1910年首次提出批判性思维概念,将其命名为“反省性思维”。“反省性思维是根据信仰或假定的知识背后的依据及可能的推论对它们进行主动、持续和缜密的思考”。格拉泽(Glaser, 1941)基于杜威的思想从三方面定义批判性思维:1)倾向于周全细致地考虑自身经验范围内的问题和对象;2)了解逻辑质询和推理的模式;3)应用这些模式的技巧。批判性思维研究代表人物恩尼斯(Ennis)认为,批判性思维是“对论点进行正确地评估”。他从12个方面将批判性思维能力要素进行分类,又将一系列倾向加入了批判性思维概念中。美国批判性思维国家高层理事会主席理查德·保罗(Richard Paul)1999年将批判性思维定义为:积极、熟练地解析、应用、分析、综合并评估支配信念和行为的有关信息的过程。综上,多数研究指出,批判性思维包括三个要素:技能、倾向和知识。其中,技能指符合逻辑探索和推理的技能;倾向指一种价值观和态度,使人倾向于持续努力参与解决问题和通过使用批判性思维评估别人的论点;知识是和学科相关的信息。

国内外学者一般从批判性思维的技能 and 倾向两方面理解批判性思维能力,并尝试利用测量工具测评批判性思维水平。综观国外学术界,批判性思维测评研究可以大体上归纳为以下三种,即多项选择测试、建构性反馈测试(constructed-responses)以及

通过在线平台收集数据评估。表一列出了三类批判性思维的测评方式,从其开放性是否独立于特定情境、是否包含批判性思维倾向、测量要素以及使用对象等方面比较几种测量方式之间的异同。

很多学者尝试翻译和修订批判性思维量表。例如,罗清旭等人(2001)翻译和修订了CCTST和CCTDI量表。彭美慈等(2004)修订了CCTDI量表,形成了CCTDI-CV量表中文版。赵婷婷等(2015)翻译和修订了美国ETS水平轮廓测试,形成了中国大学生批判性思维能力测试。

(二) 游戏化学习

在线游戏设计和实施路径聚焦于两大主轴:一是education in games,亦即利用现成的商业游戏产品,或称为commercial-off-the-shelf(现货供应,简称COTS),应用于教学或学习(Charsky & Mims, 2008)。例如,文明帝国、模拟城市。二是聚焦在games in education,即开发严肃游戏(serious games)。不同于商业游戏,其发展目的是教育或是学习用途。如今,游戏已被用作学习和评估的激励工具(Connolly et al., 2012)。游戏化指在非游戏环境中使用游戏设计元素(如挑战、分数和级别)和原则(如竞争和合作)提高用户参与度和体验(Zichermann & Cunningham, 2011)。游戏测评作为正规教育评估的补充,能够使学生更好地完成学习任务,支持高阶思维的培养(约翰·班农, 2009)。

虽然教育游戏的好处很多,但是回顾过去的研究可以发现:1)商业游戏虽然内容丰富,种类繁多,但毕竟不是以教学或学习为目的,可能存在游戏关卡过于困难,包含暴力等问题。另外,商业游戏虽然可以降低开发成本,但需要教师花时间挑选;再加上这类游戏通常不提供学生学习状态的分析,使教师难以掌握学生的学习进度与状态(赵永乐等, 2016)。2)大多数游戏化学习侧重学科知识(例如,数学、语文、科学等)的学习与成效测量,而应用于思维训练,尤其是对批判性思维测评的在线游戏相对较少。

(三) 游戏测评技术

游戏化测评是游戏化学习的一部分,测试者在测试过程中会提高相应的能力与知识。采用游戏化测评能极大地提高测评者的兴趣与动机。游戏化使用了在线或离线游戏的元素,为用户带来良好的体验,

表一 批判性思维的测评方式

测量形式	是否独立于情境	量表名称	测评方面	测量要素	适用对象
多项选择	是	《沃森-格拉瑟批判性思维量表》(简称 WGCTA)	技能	归纳推理、假设识别、演绎推理	9 年级以上学生
		《康纳尔批判性思维测验 Z》(简称 CCTT-Z)	技能	归纳、可信性、观察、预测、实验计划、谬误演绎、定义、假设确认	高中生、大学生、成人
		《康纳尔批判性思维测验, X 水平》(简称 CCTT-X)	技能	引入、推论、观察/可信度、假设	4-14 年级学生
		《加利福尼亚批判性思维技能测量量表》(简称 CCTST)	技能	解释、论点的分析与评价、演绎、归纳	大学生、高中生
		《加利福尼亚批判性思维倾向量表》(简称 CCTDI)	倾向	寻求真理性、思想开放性、分析性、系统性、自信性、好询问性、成熟性	高中生、大学生
建构性反馈测试	日常生活/有意义的情境	《恩尼斯-韦尔批判性思维作文测验》(简称 EWCTET)	技能	获得要点、发现理由和假设、产生好的推论、看出其他的可能性、避免模棱两可、发现无关性和循环论证、情绪性语言的使用等	7 年级至大学生
		《国际批判性思维测评中心批判性思维阅读与写作测试》(简称 ICTRWT)	技能	解读、分析、评价、模仿	大学生
在线平台收集评估数据	真实情境	实践性探究式模型	自发性展示技能和倾向	触发事件、探究、整合、解决方案	大学生
		纽曼(Newman)批判性思维内容分析框架		相关性、重要性、新颖性、拓展性、明确性 观点的联系、批判性评价、实际应用、理解的广度	大学生
		批判性思维模型的四种主要过程		说明、评估、推理、策略	大学生

学生不会觉得他们在考试或测试,从而提高他们的参与度、动机与交互时间(Seufert et al., 2016)。有研究表明,以故事的形式开展评估,不仅有助于评估过程,而且能让学生沉浸其中,使他们完成这些评估的速度更快(Lee et al., 2013)。马颖峰等(2019)指出,游戏化测评技术增加的游戏元素能够激发测评对象的兴趣、增加测评的趣味性及吸引力,提高用户的参与动机。

另外,传统的测评技术只能在课程学习结束后测试,无法记录学生学习过程行为,缺乏有效的反馈(Ali et al., 2013)。使用计算机评价反馈可以促进学习者学习,利用过程数据代替总结性评价,从而打破传统测评的弊端(Tanes et al., 2011; Norcini, 2010)。在线游戏评估可用于形成性评价和总结性评价;总结性评估可以采用评分或认证等任务的证据;形成性评价可通过学习者/用户使用游戏的行为和表现,如利用日志数据、用户信息轨迹评价学生行为及表现。

在线游戏可记录大量的用户行为数据,然而让数据指导教学很难。游戏化测评技术解决了这一问题。游戏化测评技术,尤其是题库型游戏,将游戏化元素和能力测量指标相结合,通过测评对象在一系

列游戏中产生的行为数据实现对其能力的评估和测评,提取隐含、具有潜在应用价值的信息(李志谢等,2019)。实时记录测评对象的行为数据,监测测评者批判性思维的行为过程,然后对这些行为数据进行分析可以告知测评者如何运用批判性思维进行决策(张青根等,2018)。因此,游戏化测评技术的出现不仅可以监控学生的学习轨迹,也可以利用海量数据提取有价值的信息。

(四) 现有批判性思维测评的瓶颈

就目前的研究看,测评工具不断更迭,批判性思维测评的设计越来越趋于多样化、学科化、动态化、具体化,然而批判性思维测评还存在一定的瓶颈。

从测评方式看,批判性思维的多项选择题量表只能衡量测评者的认知和知识水平,不能充分反映受测者的性格特点,不能揭示出受测者选择特定答案的原因和批判性思考过程(Ku, 2009)。而批判性思维能力大多是内隐加工方式,传统测评无法将内部加工机制呈现出来,无法揭示其思考过程。其次,半开放及开放性问题的建构性反馈测试,相比多项选择测试,反馈更具开放性。考官需要探究测试者解决复杂问题时是否运用了批判性思维技能。由于学习者对某一问题有不同的回答,因此,对测试结

果进行评分比较费时。指导和培训评估者也需花费较长的时间和人力。同时测试题的答案存在一定的灵活性,所以,相对于多项选择测试而言,建构性回答反馈测试评估结果缺少连续性(Norris & Ennis, 1989; Swartz & Parks, 1994)。鉴于此,研究者开始利用文本话语分析将批判性思维外显化,但是使用在线文本话语分析需要研究者对大量文本进行质性分析,在大规模测评中没有优势。

从适用对象看,针对中小学生的测评工具较少,目前国内对于批判性思维的研究主要采用汉译或国外研制的修订版批判性思维测量表或自编量表,测评工具形式单一,绝大多数自编量表缺乏大规模测试和可靠的信效度检验。因此,如何对中小学生的批判性思维进行大规模过程数据采集以及批判性思维的有效分析一直悬而未解。

三、游戏设计与开发

基于文献的梳理发现,我国对中小学批判性思维测评缺乏科学的测评工具,因此,本研究以 CCTT-X 为依据,立足游戏化测试的呈现模式,设计开发了一款基于 Web 的角色扮演类教育游戏——《Nicola Expedition(尼古拉沙漠探险记)》,它主要以 K12 阶段学生为目标用户群体,采用基于设计的研究范式并结合系统开发法,实现对测评过程和结果的记录,通过教育游戏的形式诊断学生批判性思维能力的四方面特质:归纳、推理、观察(判断可信度)、辨别假设。

(一) 游戏架构与场景设计

基于游戏化元素设计以及测评的需求,本研究设计的《Nicola Expedition(尼古拉沙漠探险记)》架构见图 1:

测试者登陆进入游戏化界面。游戏背景假设现

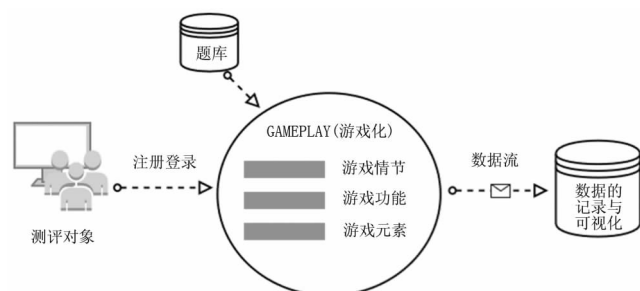


图 1 《尼古拉沙漠探险记》游戏架构

在是 2101 年 6 月 1 日,玩家作为尼古拉探险团第二团小队长,将从地球出发,准备前往一个最近才发现的星球——尼古拉。尼古拉沙漠中有个小村庄,周围是一望无际的沙漠。两年前第一团已经登陆尼古拉,到现在一直没有消息,据说以前有劫匪出没,只留下神秘的沙匪传说,所以总部派遣第二团去尼古拉,调查第一团究竟出了什么事。本游戏分五个场景,四个关键情节。游戏功能包括反馈评价、计时、提示以及实时数据记录等。

(二) 游戏交互设计

游戏中学习者可以根据自身习惯进行探索,每人可以选择不同的游戏路径,良好的交互设计能够让学习者对游戏进程有足够的控制权。在这样的游戏中,学习者自身对游戏进程的影响较大,同时,学习者本身对自我身份的认同感和价值感也会随游戏而不断增强。

本游戏的交互设计分析分为以下几个方面:首先是设计反馈激励机制。游戏测评需要设计良好的激励反馈机制鼓励学习者沉浸在游戏中,激发他们对故事内容的兴趣。在本游戏中,玩家完成某项任务后,会得到“宝箱”“钥匙”“信件恢复水”等线索奖励,激发游戏者的求知欲和好奇心。其次是设计冲突竞争和挑战机制。游戏中的冲突和竞争可以增强学习者的游戏动力,激发学习者的游戏激情,促进其在激烈的挑战中学习。在游戏情节设置中,本游戏多处设定冲突或挑战,例如,队友之间观点分歧,以考察游戏者如何做判断。最后,游戏配乐与音效能够增强测试者情景投入,提升玩家的游戏体验,引起共鸣。本游戏设计了配乐,为玩家提供良好的视听效果,采用与游戏画面风格一致的轻音乐为玩家营造轻松愉悦的氛围,如“冒险岛登录”“时间之路”“天空之城”等,并在游戏高潮时用配乐“通天塔组队任务”烘托紧张的游戏情节。本游戏还设计了“玩家跑动”“书本翻页”“宝箱打开”等互动音效,带领游戏者切身融入故事情节中,增强游戏画面感,让玩家获得生动的体验。

四、预实验

(一) 实验设计

游戏共有 49 道题,其中包括 4 道例题(1、2、19、41),均记入总分。由于故事情节影响,游戏版测试

工具测试时间不超过 50 分钟。

(二) 实验对象及过程

本研究选取上海 25 名初高中学生(初一至高三)为测试对象,年龄在 13 - 19 岁之间,实验时长 50 分钟,每个任务由学生自主控制时长,允许提前提交。最终,剔除作答题目少于 29 题的 4 人,得到有效数据 21 人。由于参与实验人数较少,所以严格来说是先导/预备实验。

五、数据结果

为了更好地展现研究结果,本文从集体表现和个人表现两方面分析数据结果。

(一) 题目难度

在本次测试的 49 道题中,当 60% 以上的测试对象答错时,我们认为这道题的答题情况较差;当答错人数介于 40% - 60% 时,我们规定这样的水平为一般;60% 以上的人数答对,说明这个题目答题情况良好。学生答题水平统计见图 2。对学生答题进行统计发现,25 道题达到较好水平,9 道题答题情况不乐观,约占 18.4%,其中 27 题和 38 题答错人数最多。这说明测试题难度适中。

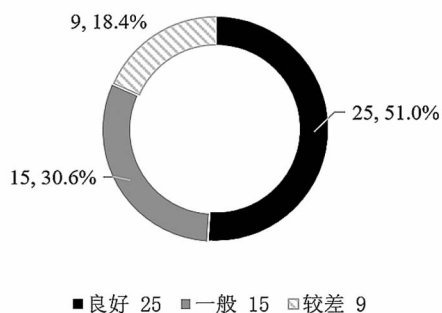


图 2 学生答题水平统计

(二) 批判性思维技能整体水平

将学生答题与批判性思维能力的四个测量维度:归纳(induction)、观察与可靠性(observation & credibility)、演绎推理(deduction)、辨别假设(assumption)对应,分别计算四个维度的回答正确率,能够判断学生批判性思维技能的整体水平。比较发现,四个测量维度题目正确率水平均在 50% 以上(见图 3),这在一定程度上反映了学生整体的批判性思维技能尚好,尤其是演绎推理和假设维度相对

较强。

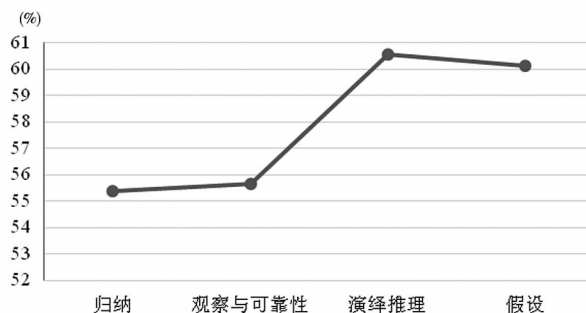


图 3 整体能力水平(均值)

(三) 批判性思维技能个体水平

本研究将每个学生的答题与批判性思维能力的四个测量维度相对应,分别计算每个学生四个维度的得分,从而判断每个学生的批判性思维技能水平(见图 4)。

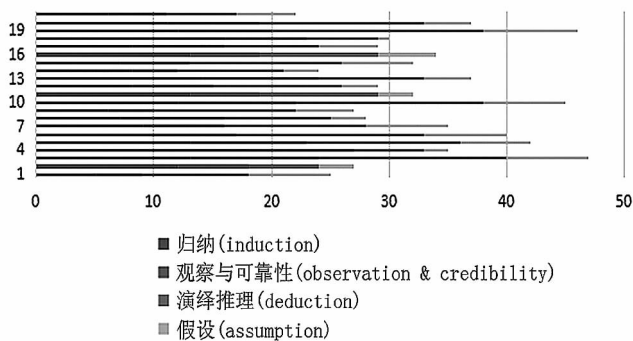


图 4 学生在批判性思维能力四个维度的得分

此外,本研究以学生个体为例对数据进行处理分析,将学生得分线与各能力总分线相对比,可以得到与整体的差距(如以学生 S1 号为例,见图 5)。我们能直观判断出该生演绎推理能力、辨别假设能力水平较高,归纳能力次之,观察(可信度)能力较弱。我们可以将结果反馈给参与测试的学生,帮助学生自我诊断。

(四) 学生查看例题与总分

游戏能记录测试者查看例题的时间,我们将查看例题的时间与总分比较,发现学生查看例题时间越多,其得分越高(见图 6)。

(五) 相关变量回归分析

为了深入探究游戏过程中涉及到的一些变量与批判性思维能力水平关系,本文选择学生性别、年

龄、查看例题时间这三个可能影响批判性思维能力(得分)变量,利用 JMP 统计分析软件,将这三者与分数变量进行回归分析。通过构造回归模型,以所得分数为因变量,以性别、年龄、查看例题时间为自变量,运用逐步法拟合回归模型,筛选出年龄、查看例题时间两个变量进入方程。结果显示,性别的 P 值概率 >0.05,这说明学生得分情况与性别不相关。进一步运行模型,进行回归方程分析及参数检验。可以分析年龄、查看例题时间与因变量“得分”的影响关系。在参数估计值报表中,年龄和查看例题时间 P 值较小 (<0.05) 这表示两个因子在预测得分情况时贡献显著。

根据以往研究结果预期学生所得分数与性别无关,且随着年龄的增长,批判性思维能力水平越高。另外,预期学生所得分数与理解例题的时间成正比。分析结果与预期结果一致,说明该游戏测评工具实证效果较为理想。

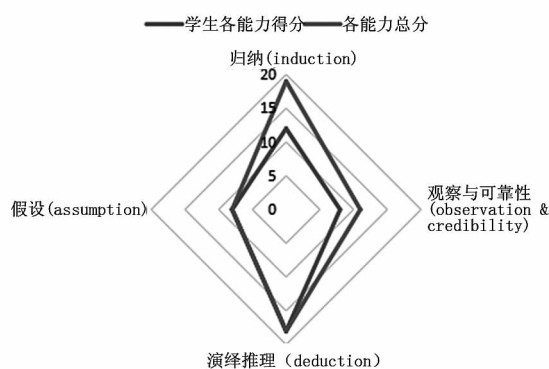


图5 学生 S1 的能力水平与整体对比

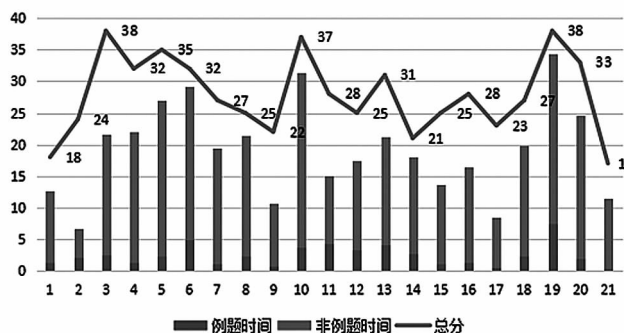


图6 查看例题时间与总分的对比

六、结论与展望

分析实验结果,我们可以得到以下结论:

(一)测评对象的批判性思维能力整体处中等水平

根据上述数据处理结果,我们得出该批次初高中学生分数均值为 27.9,81.6% 的题目答题情况较理想,且依据测试结果显示归纳、观察与可靠性的平均值占比超过各自总分的 50%,演绎推理、辨别假设的平均分达总分 60% 以上,这说明学生批判性思维能力处于中等水平。

(二)游戏化测评工具可以提高学生的学习动机及投入度

在测试完成后,笔者通过微信等线上通讯工具了解用户游戏过程的主观体验,笔者发现,该游戏测量工具存在以下优势:首先,在测试内容中加入游戏元素。玩家可以通过角色扮演进行场景、兴趣点切换,玩家在游戏任务(如寻找打开宝箱的钥匙、发现让信件显示字迹的药水等)的驱动下,提高了游戏学习动机。其次,游戏设计可以提供连续快速的结果反馈,有助于学生依据结果作出自我诊断。这与帕帕斯特基欧(Papastergiou, 2009)和赫罗德托等人(Herodotou, et al., 2015)在学科教育中采用游戏化教学方法的研究结果相同,有效营造出激励机制的教学环境,能成功激发学习者的参与度、好奇心和创造力。

(三)批判性思维得分高的学生用于查看例题的时间更多

基于游戏化测量可以实时记录用户的操作。本研究提取了用户提交答案时的时间戳,用来代表游戏者阅读和思考该题目的时间。数据处理分析发现,学生作答时查看例题的时间与得分明显相关。

本研究还存在一定的局限性:首先,学生批判性思维能力水平的诊断需要长时间观察,本研究时间条件限制,无法追踪研究对象的批判性思维能力发展情况,因此也无法验证本次测试对批判性思维能力的培养和迁移作用。其次,数据是学习分析的信息基础,现阶段游戏版本仍在更新。基于研究环境的约束,本研究只选取了少量中小学生参与游戏测试,主要是利用其初步结果优化设计游戏化测评,以便在后续研究中大范围收集数据。最后,为了更好

地验证基于题库型游戏的能力测评工具的效果,未来的研究设计可以将游戏测评和纸笔形式问卷进行对比研究。

参考文献:

- [1] Ab Kadir, M. A. (2017). What teacher knowledge matters in effectively developing critical thinkers in the 21st century curriculum? [J] *Thinking Skills and Creativity*, (23): 79-90.
- [2] Ali, L., Asadi, M., & Hatala, M. (2013). Factors influencing beliefs for adoption of a learning analytics tool: An empirical study [J]. *Computers & Education*, 62(2): 130-148.
- [3] Alnofaie, H. (2013). A framework for implementing critical thinking as a language pedagogy in EFL preparatory programmes [J]. *Thinking Skills and Creativity*, (10): 154-158.
- [4] Bassham, G., Irwin, W., Nardone, H., & Wallace, J. (2010). *Critical thinking: A student's introduction* [M]. New York, NY: McGraw-Hill.
- [5] Bedir, H. (2016). The perspectives of young adolescent EFL learners on critical thinking skills [J]. *International Online Journal of Education and Teaching (IOJET)*, 3(3): 229-238.
- [6] Charsky, D., & Mims, C. (2008). Integrating commercial off-the-shelf video games into school curriculums [J]. *TechTrends*, 52(5): 38-44.
- [7] 褚宏启(2016). 国民核心素养清单与重点 [J]. *中小学管理*, (6): 57.
- [8] Connolly, T. M., Boyle, E. A., Macarthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games [J]. *Computers and Education*, 59(2): 661-686.
- [9] Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century [J]. *Thinking Skills and Creativity*, (12): 43-52.
- [10] Facione, P. A. (2013). *Critical thinking: What it is and why it counts* [M]. Millbrae, CA: California Academic Press.
- [11] Glaser, E. M. (1941). *An experiment in the development of critical thinking* [M]. New York: Teachers College, Columbia University.
- [12] Herodotou, C., Winters, N., & Kambouri, M. (2015). An iterative, multidisciplinary approach to studying digital play motivation: The Model of Game Motivation [J]. *Games and Culture*, 10(3): 249-268.
- [13] 核心素养研究课题组(2016). 中国学生发展核心素养 [J]. *中国教育学报*, (10): 1-3.
- [14] Lee, M. J., Ko, A. J., & Kwan, I. (2013). In-Game assessments increase Novice Programmers' Engagement and Level Completion Speed [A]. *Proceedings of the ninth annual international ACM conference on International computing education research - ICER [C]*, (13): 153-160.
- [15] Li, L. (2016). Integrating thinking skills in foreign language learning: What can we learn from teachers' perspectives? [J] *Thinking Skills and Creativity*, (22): 273-288.
- [16] Ku, K. Y. L. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format [J]. *Thinking Skills and Creativity*, 4(1): 70-76.
- [17] 冷静,郭日发,侯嫣茹,顾小清(2018). 促进大学生批判性思维的在线活动设计研究及可视化分析 [J]. *电化教育研究*, 39(10): 75-82.
- [18] 罗清旭,杨鑫辉(2001). 《加利福尼亚批判性思维倾向问卷》中文版的初步修订 [J]. *心理发展与教育*, (3): 47-51.
- [19] 李志,谢思捷,赵小迪(2019). 游戏化测评技术在人才选拔中的应用 [J]. *改革*, (4): 149-159.
- [20] 马颖峰,罗晓,白羽(2015). 基于多元动机框架理论的教育游戏设计 [J]. *远程教育杂志*, 33(6): 99-105.
- [21] Miri, B., David, B. C., & Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking [J]. *Research in Science Education*, 37(4): 353-369.
- [22] 玛雅·比亚利克,查尔斯·菲德尔,洪一鸣,盛群力(2019). 21世纪的技能与元学习:学生应该学什么 [J]. *开放教育研究*, 25(1): 37-46.
- [23] Norcini, J. (2010). The power of feedback [J]. *Medical Education*, 44(1): 16-17.
- [24] Norris, S. P., & Ennis R. H. (1989). *Evaluating critical thinking* [M]. The practitioner's guide to teaching thinking series. Pacific Grove, CA: Midwest Publications.
- [25] Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation [J]. *Computers and Education*, 52(1): 1-12.
- [26] 彭美慈,汪国成,陈基乐,陈满辉,白洪海,李守国,李继平,蔡芸芳,王君俏,殷磊(2004). 批判性思维能力测量表的信效度测试研究 [J]. *中华护理杂志*, (9): 7-10.
- [27] Seufert, M., Burger, V., Lorey, K., Seith, A., Loh, F., & Tran-Gia, P. (2016). Assessment of subjective influence and trust with an online social network game [J]. *Computers in Human Behavior*, (64), 233-246.
- [28] Stuppel, E. J. N., Maratos, F. A., Elander, J., Hunt, T. E., Cheung, K. Y. F., & Aubeeluck, A. V. (2017). Development of the critical thinking tool kit (CriTT): A measure of student attitudes and beliefs about critical thinking [J]. *Thinking Skills and Creativity*, (23), 91-100.
- [29] Swartz, R. J., & Parks, S. (1994). *Infusing critical and creative thinking into content instruction: A lesson design handbook for the elementary grades* [M]. Pacific Grove, CA: Critical Thinking Press and Software.
- [30] Tanes, Z., Arnold, K. E., King, A. S., & Remnet, M. A. (2011). Using signals for appropriate feedback: Perceptions and practices [J]. *Computers and Education*, 57(4): 2414-2422.

[31] Trilling, B., & Fadel, C. (2009). 21st century skills: Learning for life in our times[M]. San Francisco, CA: Jossey-Bass.

[32] Yemenici, A. (2016). Peace education: Training for an evolved consciousness of non-violence[J]. All Azimuth, 5(1): 5-42.

[33] 约翰·班农(2009). 在线电子游戏中高阶学习技能的培养(英文)[J]. 开放教育研究,15(3):59-66.

[34] 钟启泉(2016). 基于核心素养的课程发展:挑战与课题[J]. 全球教育展望, (1):3-25.

[35] 张青根,沈红(2018). 一流大学本科生批判性思维能力水平及其增值:基于对全国83所高校本科生能力测评的实证分析[J]. 教育研究,39(12):109-117.

[36] Zichermann, G., & Cunningham, C. (2011). Gamification by design: Implementing game mechanics in web and mobile apps[M]. Sebastopol, CA: O'Reilly Media.

[37] 赵婷婷,杨翊,刘欧,毛丽阳(2015). 大学生学习成果评价的新途径:EPP(中国)批判性思维能力试测报告[J]. 教育研究, 36(9): 64-71 + 118.

[38] 赵永乐,何莹(2016). 电子游戏与教育游戏研究的困局和反思[J]. 开放教育研究, 22(4):32-39.

(编辑:赵晓丽)

Research on Assessment of Critical Thinking Ability Based on an Item Bank Game

LENG Jing^{1,2} & LU Xiaoxu²

- (1. Department of Educational Information Technology, East China Normal University, Shanghai 200062, China;
2. Shanghai Engineering Technology Research Center of Digital Education Equipment, East China Normal University, Shanghai 200062, China)

Abstract: *As a prerequisite skill for the 21st-century learners, the cultivation and assessment of critical thinking ability become increasingly important. The paper intends to break through the traditional paper-and-pencil test and design a tool to measure the critical thinking ability of primary and middle school students. In recent years, gamification has become an important way to assess learning in the era of big data, which helps provide immediate feedback to participants based on the observed learning behaviors. In this study, an item-bank game is designed to combine gamification elements with indicators of critical thinking abilities, so that participants can make corresponding decisions in game-based situations. Through online behaviors, their ability level is obtained. The results showed that the critical thinking ability of the participants was at a medium level. Those who scored higher spent more time reading important information in game situations, such as examples. Virtual game elements were used in this study to engage participants in the task situation, so as to realize the assessment of critical thinking ability in a more authentic and efficient way. At the same time, it is expected to use gamification assessment technology to further promote the development of other high-order thinking ability assessment in the big data era.*

Keyword: *gamification assessment technology; game design and development; item bank game*