

第四届全国教育实证研究优秀成果奖 获奖论文目录

评奖委员会 编

目录

优秀学术论文奖（按作者姓名笔画排序）

高校教师队伍结构与科研产出——基于世界一流大学学术排名百强中美国大学数据的分析 北京大学教育经济研究所 由由;闵维方;周慧珺 本文刊于《清华大学教育研究》2017年 第3期.....	3
中国义务教育均衡发展报告——基于《教育规划纲要》第三方评估1的证据 西南大学教 育学部 朱德全;李鹏;宋乃庆 本文刊于《华东师范大学学报(教育科学版)》2017年第1 期.....	4
大学生微信成瘾水平及影响因素调查 浙江大学教育学院数字化学习研究所 李艳;姚佳 佳;许丹莹 本文刊于《现代远程教育研究》2017年第6期.....	5
基于视频的学习过程分析:为什么?如何做? 华东师范大学课程与教学研究所 肖思汉; 德利马 本文刊于《华东师范大学学报(教育科学版)》2017年第5期.....	6
高等教育公私并举与分类管理走势分析——基于中、法、德三国经验的视角 西南财经大 学发展研究院 陈涛;邬大光 本文刊于《教育研究》2017年第7期.....	7
我国教育资源配置的区域差异缩小了吗——基于省际面板数据模型的分析 河南大学产业 经济与农村发展研究所 郑展鹏;岳帅 本文刊于《教育发展研究》2017年第9期.....	8
学校改革,价值几何——基于北京市义务教育综合改革的“学区房”溢价估计 北京大 学教育学院 哈巍;余韧哲 本文刊于《北京大学教育评论》2017年第3期.....	9
“新机制”改革对农村中小学公用经费的因果效应分析——基于准实验研究设计 南京财 经大学公共管理学院 黄斌;苗晶晶;金俊 本文刊于《中国教育学刊》2017年第11期.....	9
乡村初中班主任职业倦怠归因的个案考察——当地人的视角及其社会学的后设分析 南 京大学中华民国史研究中心江西分中心 戚务念 本文刊于《湖南师范大学教育科学学报》 2017年第6期.....	10
我国高校论文发表的变化趋势:全要素生产率的视角 北京大学教育学院 康乐;陈晓宇 本文刊于《北京大学教育评论》2018年第1期.....	11

优秀学位论文奖（按作者姓名笔画排序）

0-6岁自闭症儿童社交能力评估系统的研发与应用 于新宇 华东师范大学 硕士.....	11
美国精英高校录取决策机制研究:多重逻辑作用模型的建构 万圆 厦门大学 博士13	
初代大学生深度学习和学习结果的关系研究——基于一所研究型大学的实证案例 王帆 复旦大学 硕士.....	17
贫困地区学生高考表现与高等教育选择 刘佳辰 中国人民大学 硕士.....	19
职业教育学习评价效用的制度分析——一个链式中介模型的应用实证研究 李鹏 西 南大学 博士.....	21
技术促进小学数学深度学习的实证研究 胡航 西南大学 博士.....	26
基于行为过程表现测量合作问题解决能力的研究 袁建林 北京师范大学 博士29	
文化生态视野下藏族中学理科教育问题研究 党宝宝 西北师范大学 博士.....	33
被默许的误认——当代大学教师对教育者身份的理解与建构的质性研究 曾妮 北 京大学 博士.....	36
课堂教学公平及其指标体系研究 雷晓庆 南京师范大学 硕士.....	39

高校教师队伍结构与科研产出——基于世界一流大学学术 排名百强中美国大学数据的分析

北京大学教育经济研究所 由由;闵维方;周慧珺

本文刊于《清华大学教育研究》2017年第3期

摘要: 本文基于世界一流大学学术排名数据和46所美国研究型高校10年的面板数据,分析了高校科研产出的影响因素以及高校教师队伍结构变化可能带来的影响。研究发现,高校科研产出除了由外部经济环境、高校类型、人员规模、工资水平等因素决定外,还可能受高校教师队伍的岗位、合同、职称、性别/种族结构的影响。其中,全职教师规模一定的情况下,辅助人员的增加有利于高校科研产出数量与质量的提升,非终身教职体系教师的聘用并未如预期那样有助于提高高校科研产出,助理教授得利于能够在科研上集中精力而产出较高,传统弱势群体教师比例的变化并未系统性地影响高校产出。基于上述研究发现,建议我国高校在人事制度改革中致力于发挥教师资源的比较优势,为教师提供充分的辅助支持,避免造成人力资源的浪费。

关键词: 研究型大学;科研产出;教师队伍结构;美国;高校人事制度改革

Abstract: Based on 10-year's institutional data of 46 US research universities and their ARWU rankings, this paper examines factors that influence university research performance and explores whether universities' changes in faculty structure have impacts on their research performance. This research finds that, in addition to factors such as external economic environment, institutional control, numbers of faculty and student, and salary, other university's faculty structures such as faculty's position, contract type, academic rank, and demographics are also potential determinants to influence university research performance. Specifically, given that full-time faculty numbers are certain, increase in supporting staff is proved to help enhance university research performance in quantity and in quality. However, increase in non-tenure track faculty is not suggested to be necessarily beneficial. In addition, assistant professors have relatively higher research performance, which can be partially attributed to universities' enabling environments and supports. Finally, no obvious systematic relationship is found between different demographic groups. As a new wave of personnel reform in higher education is going on in China, this paper suggests China's universities should dedicate efforts to utilizing the comparative advantages of existing faculty resources, providing them sufficient supports, and avoiding wasting human resources.

Keyword: research universities; research performance; faculty structure; the United States; personnel reform in higher education;

中国义务教育均衡发展报告——基于《教育规划纲要》第三方评估 1 的证据

西南大学教育学部暨中国基础教育质量监测协同创新中心西南大学分中心
朱德全;李鹏;宋乃庆

本文刊于《华东师范大学学报(教育科学版)》2017年第1期

摘要: 通过对苏、鲁、鄂、湘、川、渝等 14 个省(市、自治区)82 个县(区)546 所中小学的实证调查,结合 2010 年-2014 年义务教育发展的宏观数据和部分省(市、自治区)的典型案例分析,对 2010 年以来中国义务教育均衡发展状况进行监测评估。评估发现:5 年来,中国义务教育均衡发展成效显著,不仅从战略任务的高度"建立健全义务教育均衡发展保障机制",而且不断"完善投入机制"、"加大教育投入",深入推进学校标准化建设成效显著,学校办学条件大幅改善,同时,"加强师资队伍"有所进步,"均衡配置师资"逐步实现。但也存在着义务教育经费投入"中部塌陷"、办学条件"内涵化"差距巨大、农村师资队伍结构性缺编严重、城镇化进程中新挑战不断、教育质量城乡差距甚大等问题。因此,新时期中国义务教育均衡发展要坚持五大发展理念,转变义务教育发展方式,深化义务教育综合改革,开拓义务教育内涵均衡、优质均衡与特色均衡的发展道路。

关键词: 义务教育;均衡发展;第三方评估;治理体系;综合改革

Abstract: This article is based on the empirical investigation in 546 elementary and middle schools from 82 counties 14 provinces, as well as macro data of compulsory education development and case studies of each province from 2010 to 2014, It aims to evaluate the balanced development of compulsory education in China since the implementation of National Medium and Long-term Plan for Education Reform and Development(2010-2020) and presents an analysis of the balanced development of compulsory education since 2010. According to the evaluation, compulsory education in China has made remarkable achievements in terms of balanced development, reaching a new starting point. First, a guarantee mechanism has been established to promote the compulsory education in a strategic way. Second, an investment mechanism has been improved to ensure a growing investment in compulsory education. Third, remarkable achievements have been made in the standardization of school construction. Fourth, some progress has been made in capacity-building of the teaching staff, and the balanced allocation of the teaching staff is under way. However, the evaluation reveals some problems, such as inadequate funding of compulsory education in central areas, big gaps in education quality between schools, structured short-staffed schools in rural areas, plus the new challenges in urbanization. Therefore, to promote high-quality and balanced development of compulsory education in China, it's essential to adhere to the development concept of innovation, coordination, green, openness and sharing, adopting a top-level design and overall planning, and ensuring there is a cushion in place for those in great need. Besides, it's

important to build a development network, focusing on the key points to promote balanced and improved quality. First, it's important to renew the development concept and change the traditional ways of balanced development of compulsory education. Second, it's necessary to establish a linking mechanism, strengthen the responsibilities of the government to promote the balanced development and synergetic governance of compulsory education. Third, it's imperative to further improve comprehensive reform, solve the problems one by one, and overcome the major obstacles to the balanced development of compulsory education. Fourth, it's advisable to enhance the guarantee system, and explore a three-dimensional path to balanced development of compulsory education.

Keyword: compulsory education; balanced education; third-party evaluation; governance system; comprehensive reform;

大学生微信成瘾水平及影响因素调查

浙江大学教育学院数字化学习研究所 李艳;姚佳佳;许丹莹;

本文刊于《现代远程教育研究》2017年第6期

摘要: 微信作为连接人与人、人与服务、人与商业的社交媒体,已极大程度地深入并影响着使用者的工作、学习和生活。作为微信使用的重要主力军,大学生使用微信呈现出什么样的特点?成瘾水平是否较高?哪些因素影响其微信成瘾水平?从981位大学生的随机调查中发现:大学生的微信成瘾水平整体并不高,但不同性别、专业背景、受教育阶段的大学生微信成瘾水平有显著差异,且微信使用时间越久,好友以及参与微信群、关注公众号数量越多,微信成瘾度越高。另外,大学生微信阅读行为对微信成瘾水平无显著的直接影响,但它能通过交互行为和创作行为间接影响微信成瘾水平;交互行为对微信成瘾有显著负影响,创作行为对微信成瘾有显著正影响。预防大学生微信成瘾,可以通过高校学习共同体建设、学习空间改造、社团/俱乐部活动创新等方式,提高学生参与校内外/课内外活动的机会,减少其对手机和微信的依赖与成瘾可能;通过高校课堂教学设计的改进引导大学生优化使用微信的意识、习惯和能力等。

关键词: 大学生;微信;社交媒体;成瘾水平;影响因素

Abstract: As a social medium that connects people, services and business, WeChat has deeply influenced the users' work, study and life. As an important group of We Chat users, college students might have their own features of using We Chat and what the features would be? Do they have a high level of We Chat addiction? Which factors might impact their We Chat addiction level? Based on a random survey from 981 college students, the study found that: the overall level of college students' We Chat addiction was not high, however, students with different genders, majors and academic levels had significant differences

in their We Chat addiction levels; the more time college students spent on We Chat, the more friends they had, the more groups they joined, and the more public account they follow, the higher level of We Chat addiction they would have. In addition, college students' We Chat reading behavior had no significant direct effect on their We Chat addiction level, while it could have indirect effect on We Chat addiction through We Chat interactive and creative behaviors; interactive behavior had a significant negative impact on college students' We Chat addiction level, while creative behavior had a significant positive impact on their We Chat addiction. To prevent college students from We Chat addiction, universities might pay more attention to the establishment of college students' learning communities, learning space reform, innovation of community/club activities and so on, and try to improve students' participation in on-campus and off-campus activities in order to reduce the possibility of students' dependence and addiction on the mobile phone and the We Chat. At the same time, instructors might improve their instructional design to guide students to optimize their awareness, habits and ability of WeChat usage in learning.

Keyword: College Students; WeChat; Social Media; Addiction Level; Influential Factors

基于视频的学习过程分析:为什么?如何做?

华东师范大学课程与教学研究所 肖思汉;德利马;

本文刊于《华东师范大学学报(教育科学版)》2017年第5期

摘要: 经过 40 多年的发展,基于视频的实证研究已经成为教育研究——尤其是对学习过程的研究——的主流取径之一。然而,对于为什么要运用视频,如何对视频进行分析,学界仍然见解不一。本文在梳理相关历史和理论的基础上,通过两个案例呈现视频分析的三个重要维度,即个人的身体、社会的身体、身体与物件的互动。本文指出,这三个维度根本性地影响着我们对课堂互动、学习过程、教学机制的分析、理解与阐释,而视频数据在其中起到了不可或缺的作用。最后,本文强调,为了发掘视频研究的潜力,教育工作者需要从理论视角上扭转“语言中心论”的偏见,探索教师对自己课堂的评价权,并且尊重课堂生活世界中的“日用行常”。

关键词: 课堂视频;课堂话语;课堂互动;视频分析;互动分析

Abstract: After more than four decades of development, video-based analysis has become a major approach in educational research, in particular analysis of learning processes. Yet, with regard to the “why” and the “how-to,” video researchers have not achieved a consensus. Based on a brief review of history and relevant theories, this article uses two case studies to illustrate three crucial aspects of video-based analysis: the individual body, the social bodies, and the

bodies-things interaction. We argue that these aspects profoundly shape our analyses and understandings of classroom interaction, learning processes, and mechanism of teaching, and that video data play a critical role therein. Finally, we argue that educators and educational researchers need to avoid the bias of logocentrism, to empower teachers to evaluate their own classroom, and to respect the everyday workings of classroom life-worlds, if we are to fully appreciate the potential of video-based analysis.

Keyword: classroom video; classroom discourse; classroom interaction; video-based analysis; interaction analysis

高等教育公私并举与分类管理走势分析——基于中、法、德三国经验的视角

西南财经大学发展研究院 陈涛;邬大光

本文刊于《教育研究》2017年第7期

摘要: 二战之后,世界各国高等教育的举办主体发生变化,泾渭分明的公私界限逐渐模糊,高等教育办学体制从"国家举办"到"公私并举",高等教育管理体制从"高度集中"到"权力下放",高等教育投资体制从"政府拨款"到"多元筹资"。传统意义上的公立、私立高校的分类遇到了困境,判别维度不断"升级",区分标准更加复杂化,对各国高等教育制定新的管理政策提出了挑战,突破公私二元的分类管理模式成为中、法、德三国应对政策挑战的主要手段。我国民办高等教育分类管理政策应坚持资金分配的公共性原则、治理模式的平衡性原则、质量评估的自主性原则、公平发展的协调性原则。

关键词: 高等教育;公私并举;分类管理

Abstract: After the Second World War, as changes have taken place in the world-wide running bodies of higher education, the boundary between public and private that was once clear is becoming blurry. School-running system of higher education varied from "state-running" to "simultaneous development of the public and private", management system of higher education varied from "highly centralized" to "decentralized", and investment system of higher education varied from "government appropriation" to "multi-channel financing". Traditionally, the public and private colleges and universities face the difficulties in classification. Since the dimensions of classification are increasingly "updated" and the distinguishing standards are becoming more complicated, challenges have been put forward on new policy-making on higher education in many countries. The classification management model that breaks the duality of the public and private has become the primary tool to cope with the policy challenges in China, France and Germany. In China, the policy for

classification management of private higher education should stick to the public principle of allocation of funds, the balanced principle of governance model, the autonomy principle of quality evaluation, and the coordinative principle of fair development.

Keyword: higher education; simultaneous development of the public and private; classified management

我国教育资源配置的区域差异缩小了吗——基于省际面板数据模型的分析

河南大学产业经济与农村发展研究所 郑展鹏;岳帅

本文刊于《教育发展研究》2017年第9期

摘要: 本文利用新古典经济增长理论的收敛假说,基于省际面板数据模型,对我国各教育阶段教育资源配置区域差异的动态演变趋势进行了实证分析。结果显示:我国不同教育阶段教育资源配置区域差异的动态演变趋势不同,具体来说,就 σ 变异系数而言,义务教育和高等教育阶段的教育资源配置区域差异均呈现缩小的趋势,而中等教育阶段区域差异未表现出缩小的动态演变特征。从 β 绝对收敛来看,全国及三大俱乐部在义务教育阶段教育资源配置的区域差异均不断缩小,中等教育阶段不断扩大,高等教育阶段没有呈现出明显的动态演变特征。从 β 条件收敛来看,全国及三大俱乐部各教育阶段教育资源配置的区域差异均呈现出不断缩小趋势,且与 β 绝对收敛相比,区域差异缩小的速度有所加快,但各教育阶段区域差异缩小的速度不同,高等教育阶段最快,义务教育阶段次之,中等教育阶段最慢。同时,经济发展水平、财政分权、对外开放度及人口老龄化对不同地区各教育阶段教育资源配置的影响也有所不同。

关键词: 教育资源配置;区域差异;收敛性

Abstract: Based on the provincial panel data model, this paper makes an empirical analysis on the convergence of regional education resources allocation in china. The results showed that different configuration education resources in our country the convergence of different, specifically, compulsory education and higher education resource allocation showed the σ convergence trend, and the secondary education stage did not show the characteristics of σ convergence. From the point of The β absolute convergence view, and the three big clubs in the compulsory education stage shows β absolute convergence, the secondary education stage showed the trend of dynamic evolution of divergence, the stage of higher education showed no obvious divergence. From the view of the β conditional convergence, and the three clubs in each education stage shows β convergence, but the convergence rate of different stages is different higher education, the fastest convergence speed, the second fastest is

compulsory education, and secondary education has the slowest convergence rate. At the same time, the influence of the level of economic development, fiscal decentralization, open to the outside world and the aging of the population on the allocation of educational resources in different areas of education.

Keyword: education resources allocation; area differentiation; convergence

学校改革,价值几何——基于北京市义务教育综合改革的“学区房”溢价估计

北京大学教育学院 哈巍;余韧哲

本文刊于《北京大学教育评论》2017年第3期

摘要: 近年来国内关于学校质量资本化的研究不断涌现,但这些研究都未能很好地解决学校质量与房价之间的内生性。本文结合了北京市义务教育综合改革带来的外生政策冲击和边界断点回归这两种策略,估计了“学区房”的溢价,考察了薄弱小学质量提升的货币价值。研究发现,对于2014-2016年间改革的薄弱学校,其服务范围内的房价平均只提升1.2%,且这一提升效应在学校宣布改革至少一年之后才得以显现。但是对于第一批实施改革的小学 and 60平方米以下的小户型房屋而言,这种价格传导机制更加明显而且更加迅速。这意味着此次改革带来的优质教育资源可能被相对优势群体获得,这对未来我国义务教育阶段均衡化导向政策的制定具有一定启示。

关键词: 学区房;公共品资本化;义务教育综合改革;事件研究模型;北京

“新机制”改革对农村中小学公用经费的因果效应分析—— 基于准实验研究设计

南京财经大学公共管理学院 黄斌;苗晶晶;金俊;

本文刊于《中国教育学刊》2017年第11期

摘要: 利用全国县级数据,采用倾向得分匹配结合倍差法,就“新机制”改革对农村中小学校公用经费的水平效应与分配效应进行因果分析。结果显示,改革具有较强的水平效应,使农村中小學生均公用经费水平大幅度上升。相比之下,改革的分配效应差强人意:一方面,改革对于城乡中小學生均公用经费差距不具有显著的改善影响,甚至有证据显示它可能扩大了此差距;另一方面,改革虽能缩小东中西部

小学生均公用经费差距,但对初中生均公用经费地区间差距不具有改善作用。为此,建议突破现有保障型体制,构建能充分发挥分配效应的新体制;公用经费支出职责上移,人员经费多级政府共担;打破城乡分治模式,进一步推动义务教育城乡一体化改革。

关键词: “新机制”改革;农村学校;公用经费;因果效应

乡村初中班主任职业倦怠归因的个案考察——当地人的视角及其社会学的后设分析

南京大学中华民国史研究中心江西分中心 戚务念;

本文刊于《湖南师范大学教育科学学报》2017年第6期

摘要: 乡村学校班主任的职业倦怠,并不仅仅是个心理学问题,而是有其特殊的社会学根源。根据在一所乡镇农村初中的实地调研发现,从当地人的视角分析班主任的职业倦怠原因,主要有经济收入的参照对比、管理制度的设计缺陷、乡村教育的衰败、教师职业高危险性、岗位胜任力不足等问题。对当地人的视角进行社会学的后设分析,乡村初中班主任的职业倦怠,可以看作是在包括时间与空间在内的社会结构性要素以及包括本领恐慌与情感枯竭恐慌在内的个体行动性要素双层夹击下形成的。对于职业倦怠这一社会现象需要进行更深层次的理论反思,如可能产生的结构性陷阱与需要层次升级链断裂等问题。从实践层面缓解这一现象,相应的政策举措需要更具系统性、整体性。

关键词: 乡村教育;职业倦怠;班主任;当地人视角;结构—行动

Abstract: The staff burnout of head teachers in rural schools is not only a psychological problem, but also has its special social source. According to the field investigation of one rural middle school, the local people attributed the burnout to the reference of economic income, the design defects of the management system, the decline of rural education, the high risk of teacher's vocation, and not qualified for the position. From the perspective of sociology, the staff burnout can be seen as a form of the social structural elements (including time and space) and the individual action (including the panic to ability and emotional exhaustion). The social phenomenon of staff burnout requires a deeper theoretical reflection, such as the possible structural pitfalls and the chain fracture of the need to upgrade. In order to alleviate this phenomenon from the practical level, the corresponding policy measures need to be more systematic and integrated.

Keyword: country education; staff burnout; head teachers; local perspective; structure-action

我国高校论文发表的变化趋势:全要素生产率的视角

北京大学教育学院/教育经济研究所 康乐;陈晓宇

本文刊于《北京大学教育评论》2018年第1期

摘要: 本文在教育部出版的《全国高校社科统计资料汇编》、《高等学校科技统计资料汇编》的基础上,结合科学网提供的SCI扩展版、SSCI和A&HCI数据,使用索洛残差法估算了1991-2012年间我国高校论文发表的全要素生产率,并分析其分布特征、变化趋势以及对我国高校论文增长的贡献。研究发现:首先,在我国高等院校的英文发表、特别是英文的科技发表中,资本投入的产出弹性更大;第二,我国"985"大学的效率优势体现在英文发表而不是中文发表;第三,我国高等院校的中文发表效率开始下降,但英文发表的效率正逐渐提高;最后,与英文科技发表相比,我国高校论文发表的效率在英文社科发表中增长更快、对增长的贡献率也更大一些。

关键词: 高等院校;论文发表;全要素生产率;索洛残差法

0-6岁自闭症儿童社交能力评估系统的研发与应用

于新宇 华东师范大学 硕士

摘要: 社会交往障碍是自闭症的核心缺陷之一,严重影响自闭症患者融入学校和社会。大量研究表明,早期干预(6岁前)对自闭症患者的行为改善和能力提升有重要意义。而早期干预是否有效,除采取的干预方法外,还依赖于精准、有效的评估。国内现有的相关评估工具较少,且存在着对评估者专业性要求高,评估过程复杂,评估结果与训练脱离等诸多问题。因此,本研究面向0-6岁自闭症儿童,以指导训练为主要目的,研发了一整套社交能力评估系统,并在人形机器人干预中进行了应用。主要研究内容如下:

1. 自闭症社交能力评估系统的内容设计

社交评估工具设计。对中残联《孤独症儿童发展评估表(社交领域)》进行修订,将现场评估形式简化为问卷形式,经检验:两者具有很好的一致性($r_s=0.848$, $p<0.01$);对原量表结果分析方式进行调整,通过对223份问卷数据进行分析,结合文献资料,将评估项目依难度进行重新排序,根据评估项目难度等信息设计社交能力定位图,用于分析儿童评估结果,制定个别化评估报告。方案制定功能设计。针对每个评估项目,设计与之相匹配的康复建议,利用康复建议,制定个别化训练方案。效果监控功能设计。将社交能力分为六个可独立监控的指标,采用三维立体图的形式,动态反映儿童社交能力的变化趋势,便于在训练过程中根据儿童的不同表现对训练方案进行调整。

2. 自闭症社交能力评估系统的程序开发

评估系统依功能可划分为社交能力评估模块、训练方案制定模块、训练效果监控模块,本研究将通过软件程序开发依次完成三大模块的功能实现。本研究主要采

用LabVIEW虚拟仪器程序设计软件作为编程工具，通过LabVIEW编程，实现评估信息输入、数据处理、结果输出等一系列过程的自动化。在此基础上，在训练方案制定模块，采用人形机器人技术，设计一整套人形机器人训练方案，目的是提供参考训练方案。评估系统最终将打包为一个完整的电脑端应用程序，并生成Windows操作系统下的安装文件，可以在任意一台装有Windows操作系统的电脑中安装使用。

3. 自闭症社交能力评估系统的实际应用
通过人形机器人干预方法，对评估系统进行应用。结果显示：（1）接受训练的10名儿童经训练成绩有明显提升（ $Z=-2.530$, $p<0.01$ ）；（2）对一名儿童进行监控数据分析，得到其在总成绩提升的基础上，各个监控指标的变化也呈现出上升趋势；（3）对一名儿童的“同伴交往行为”发生次数进行观测，得到经训练儿童的“同伴交往行为”发生次数增加。综上，由于训练的内容均来自于评估结果，因此，训练结果的有效性亦证明了评估的有效性。

本研究的创新之处体现在：1. 简化原有评估方式。对已广泛使用的评估表进行调整与修订，在保留原始评估项目的基础上缩短了评估时间，从开始评估到拿到报告仅需10分钟；2. 社交能力精准定位。设计了“社交能力定位图”，能对0-6岁自闭症儿童社交能力现状进行的精准定位与分析；3. 开发评估系统程序。利用LabVIEW编程技术进行软件编程，完成了评估系统软件程序的开发，程序易推广，有助于扩大受益对象；4. 开发效果监控系统。从评估到训练全程跟踪监控，收集儿童一整段评估数据，有利于对个案进行长期跟踪，调整方案；5. 提供评估干预一体化方案。从评估到干预提供一整套解决方案，主要工作量由计算机和机器人承担，降低了对实施人员专业性的依赖，有利于在各个地区的学校、机构、家庭中进行推广。本论文研究成果已申请三项国家发明专利（申请号：201810156670.0；201711166859_X；201711124128.9），目前进入实质审查阶段。

关键词： 自闭症谱系障碍； 社交能力； 评估系统； LabVIEW； 人形机器人

Abstract: Problems in social communication and social interaction are the types of symptoms of Autism Spectrum Disorder(ASD), which seriously affects the integration of autism patients into schools and society. A large number of studies have shown that early intervention (before age 6) is of great significance for improving behavior and enhancing abilities of autism patients, however, whether early intervention is effective or not depends on accurate and effective assessments. There are few related assessment tools available in China, and there are many problems exist. Therefore, this study is based on the LabVIEW, designs and develops an assessment system for social skills of children (0–6 Years) with ASD, with guiding training as the main purpose. The main research content is as follows:

1. Content design of the assessment system

Firstly, the study revise *the autistic children development assessment* that published by the China Disabled Persons Federation Rehabilitation Department, simplifies the original assessment method by using questionnaire instead of on-site assessments, through analysis and test, it is proved that the questionnaire has good consistency($r_s=0.848$, $p<0.01$); on this basis, the results analysis method has been adjusted and revised. Secondly, the study designs a social skills map and provide each child with a personalized training program based on the map. Finally, this study designed a training effect monitoring function, it is easy to modify and adjust the

training program according to children's different performance during the training process.

2. Programming development of the assessment system

The assessment system can be divided into three major modules: social assessment module, training program development module, training effect monitoring module according to functions. This study will implement the functions of the three modules through programming. This study mainly uses the LabVIEW software as a programming tool to automate a series of assessment processes. On this basis, in the training program development module, using humanoid robot technology to provide a reference training program. The assessment system will eventually be packaged as a complete computer-side application and generate installation files for the Windows operating system that can be installed on any computer with a Windows operating system.

3. Practical application of the assessment system

Humanoid robot intervention method is applied to the assessment system. The results show that: (1) After the training, 10 children have significantly improved their training results ($Z=-2.530, p<0.01$); (2) Analysis of a child's monitoring data, the data shows that the change of each monitoring indicator also showed an upward trend; (3) Analysis of a child's "peer interactions" occurrences, the data shows an increase in "peer interactions". In summary, because the training content comes from the assessment results, the effectiveness of the training results also proves the effectiveness of the assessment.

The **innovations**: 1. Simplify the original assessment method. Shortening the assessment time and saving assessment resources; 2. Accurate positioning of social skills. A "social competence location map" is designed to accurately locate and analyze the social status of the children; 3. Program development of assessment system. The use of LabVIEW programming technology for software programming, completed the development of evaluation system software program; 4. Case tracking and monitoring. Helps to accurately analyze cases; 5. Provide assessment intervention integration plan. The assessment and training are closely combined, facilitate the promotion of the system in schools, institutions, and families in various regions. The research results of this thesis have applied for three national invention patents (application number: 201810156670.0; 201711166859_X; 201711124128.9).

Keyword: *Autism Spectrum Disorder; Social Skills; Assessment System; LabVIEW; Humanoid Robot*

美国精英高校录取决策机制研究：多重逻辑作用模型的建构

万圆 厦门大学 博士

摘要: 美国精英高校对生源的选拔，看似简单、实则复杂，唯有获得完整、深入

地认识，才能取其精华。本研究以“录取决策机制”为切入点，基于对美国精英高校选拔生源时使用的考量因素、审阅方式以及质量控制手段的考察，分析其录取决策的目标及达成方式，进而构建驱动录取决策行为的多重逻辑作用模型。通过以美国三所精英公立大学、两所精英私立大学、两所精英文理学院为关键案例，运用质性取向下的多个案比较法进行资料的收集和分析，基于对 52 位招办人员及利益相关者的访谈和大量的相关文本，本文得出以下研究发现、结论和建议：

首先，美国精英高校均遵循目标导向、服务自身利益和使命的原则来选拔生源。一方面，美国精英高校的共同招生目标为以公正、透明的方式，塑造一届卓越且多样化的新生。另一方面，录取决策的铁律为服务高校利益和使命：高校利益包括达到注册目标、满足财政需求、提高声望、增加选拔性、促进种族多样化、保持校园体育实力等等，但各校录取决策服务的利益诉求和优先项不尽相同；录取决策实践受到高校承担的培养公民和公民领导者以及服务社会三大办学使命的驱动。

其次，为了达成各项招生目标，案例高校在录取决策中都采用了基于多个标准的综合评价和兼顾公平与效率的整体性审阅，并通过质量检查技术、组建专业审阅团队、投入大量资源、接受教师和董事会监督等手段进行质量控制。其中，在综合评价上，各校通过查看学业表现、个人成就和个人背景三个维度下的多个具体因素，基于三个标准——在大学取得成功的潜能、做出外在贡献的潜能以及从大学教育中获益的潜能，来判断申请者与高校的匹配程度。综合评价中的决策规则体现为：学业成就是几乎所有申请者参与录取竞争的必要条件，其中学业杰出者基本都会被录取；个人成就有助于区分属于中间群体的“不错的学生”；加号因素有助于学业表现和个人成就相似的中间群体和“合格的学生”赢得竞争，也可使得“不合格的学生”以降低学业标准的形式被录取。技术促进小学数学深度学习的实证研究在整体性审阅上，各校均使用背景化审阅、个体化审阅和集体化审阅，保证录取决策的公平与效率：1) 将每位申请者取得的成就置于高中教育背景和个人成长背景中考察，考虑其拥有的资源和机会，同时被置于申请者库中考察，即各校将每位申请者与来自同一高中的申请者、与平均水平以及与录取标准之间进行比较，而非脱离背景进行横向比较。2) 个体化审阅将每位申请者视为“一个整体”和“一个个体”，没有阈值、没有公式、没有单一的决胜因素，是最高法院批准的考虑种族的合法方式和州禁令下达成种族多样化的必要方式，也是保证决策自由度、区分申请者和达成各项招生目标的最佳方式。3) 集体化审阅的模式分为委员会讨论模式和团队审阅模式，但产出一届新生的路径均始于单名读者对每份申请材料进行基于量化评级的分项评价或基于质性评语的整体评价，并经历至少两名读者的评价以及委员会讨论或高级 AO 的审阅，止于确定一届新生名单的招生季末审核，是一种高度依赖人为经验的协作式决策路径。

第三，在不同类型精英高校的录取决策实践中，学术逻辑、文化逻辑、市场逻辑、政治逻辑、伦理逻辑、成本逻辑的影响同时存在，但每项逻辑发挥的作用强度存在差异：1) 学术逻辑和文化逻辑是导致各校在使用考量因素时存在相似性的逻辑，市场逻辑、政治逻辑、伦理逻辑、成本逻辑则是导致差异性的逻辑。2) 学术逻辑和伦理逻辑的作用强度均在公、私立维度存在细微差别，并均在公立大学体现更为明显。文化逻辑的作用强度在公立大学、私立大学和文理学院之间存在细微差别，其中在文理学院体现最为明显。市场逻辑和政治逻辑的作用强度在公、私立维度存在显著差别，其中市场逻辑在私立高校体现尤为明显，政治逻辑则在

公立大学体现尤为明显。成本逻辑的作用强度不存在办学类型的差异。最后，美国经验为完善我国高校招考制度提供的启思包括：拓宽卓越生源的定义，探索生源多样化的价值；促进入学机会公平，提升招生透明度；坚定推进综合评价改革，完善实践操作；扩大高校招生自主权，提供环境支持。

关键词：美国；精英高校；录取决策；整体性审阅；多重逻辑

Abstract: The selection of applicants for admission to American elite colleges and universities seems simple, but complicated indeed. It requires an all-round and in depth research to get a good understanding of it. This research adopts “decision making mechanism” as a joint point, analyzes the goals of decision-making practices and the ways of their realization based on admission factors, review approaches and methods of quality control adopted by American elite colleges and universities in undergraduate admission, and then proposes a multi-logic function model which drives admission decision-making behavior. With data collection and analysis through comparative case studies involving three elite public universities, two elite private universities and two elite liberal arts colleges in a qualitative approach, and with interviews of 52 admission officers and stakeholders as well as a large amount of relevant documents, this research has made findings and reached conclusions and suggestions as follows:

First of all, the principle of being goal-oriented, institutional interests-served and institutional mission-served is commonly abided by American elite colleges and universities in admission decision-making practices. On one way, their common goal is to craft an excellent and diversified class in a fair and transparent way.

On another, the iron law of admission decision-making is to serve the interests and mission of the institutions. Those interests include reaching enrollment goals, meeting financial needs, enhancing institutional prestige, improving institutional selectivity, promoting race diversity, maintaining sports strength, and so on. However, the specific interests and priorities in admission decision-making are different. Admission decision-making process is driven by the three missions of cultivating citizens, cultivating leaders of citizens, and serving the society. Next, in order to achieve various admission goals, the colleges and universities involved in the case studies not only have conducted multiple-criteria based comprehensive assessment and holistic review characterized by equity and efficiency, but also have applied quality check techniques, built teams of expert readers, invested lot of resources, accepted inspections from faculty for quality control. Regarding comprehensive assessment, those universities or colleges judge the fitness of applicants on the basis of the three criteria of the potential of achieving success in college, potential of making external contribution and potential of benefiting from college education, through reviewing various factors under the three dimensions of academic performance, personal achievements and personal background. The decision rules in comprehensive assessment are: academic achievement is an essential requirement, those with excellent academic performance being generally admitted; personal achievements help distinguish those “good students” in the middle; plus factors help those “good

students” in the middle and those “qualified students” win in the competition, also makes it possible for “unqualified students” to get admitted due to lowered academic standards. With regard to holistic review, the institutions in case studies all have applied contextualized review, individualized review and collective review in order to ensure the equity and efficiency in admission decision-making. First, each applicant would be reviewed under their high school educational background and personal family background, considering the availability of resources and opportunities. Meanwhile, each applicant would be reviewed in the applicant pool, which means a certain applicant would be compared with other applicants from the same high school, with average achievement, and with admission standards, rather than horizontally comparing two applicants without context. Second, individualized review regards each applicant as both “a whole” and “an individual”, with no cutoff, no formula, no single tie-breaking factor. It is a legal practice approved by the Supreme Court considering races, an essential way of realizing race diversity under the state ban of Affirmative Action, and is also the best way to ensure decision-making freedom, distinguish applicants and reaching various admission goals. Third, although there are two modes of implementing collective review, namely committee discussion mode and team reading mode, but the process of forming a new class through collective review both start with a single reader doing an analytical assessment of each application based on numerical rating or a holistic assessment based on qualitative notes, then goes through two readers’ assessment and committee discussion or there view of senior admission officers, and ends with the end-of-season review which determines the name list of the new class. Therefore, the collective review is a collaborative decision-making approach which highly relies on human judgment. Third, in the admission decision-making practices of different types of elite colleges or universities, the influences of academic logic, culture logic, marketing logic, political logic, ethical logic and cost logic co-exist, but each differentiating in their degree of effect: first, academic logic and culture logic are the logics that lead to similarities that institutions apply in the use of various factors, while marketing logic, political logic, ethical logic and cost logic are logics that lead to differences; second, the influence of academic logic and ethical logic differs in public and private higher educational institutions, and manifests more obvious in public universities. The effect of culture logic differs subtly among public universities, private universities and liberal arts colleges, while manifesting the most obvious in liberal arts colleges. The effect of marketing logic and cost logic varies a lot in public institutions and privations, with marketing logic manifesting most obvious in private universities and political logic in public universities. The effect of cost logic does not differ among different types of universities. In the end, the enlightenments of American experience for China’s college admission policies include: first, broaden the definition of excellent students and explore the value of diversification of students; second, promote equal access and enhance decision-making transparency; third, firmly propel the reform of comprehensive assessment and improve real practice; fourth, expand admission autonomy of universities and provide environment support.

Keyword: the United States; elite colleges and universities; admission

初代大学生深度学习和学习结果的关系研究——基于一所 研究型大学的实证案例

王帆 复旦大学 硕士

摘要：新兴科技急速发展隐然有替代劳动力的趋势，高校和用人单位愈加重视学生的深层次能力，深度学习方式对高等教育质量有重要影响，业已成为本科学习研究的热点之一。学者大多发现深度学习与学生能力获得存在显著正向相关关系（Laird et al., 2014；史秋衡，郭建鹏，2012），与学生深造期望存在正向相关关系（Rocconi, Ribera, & Laird, 2015），而深度学习对学业成绩影响，研究尚存争论。同时，高等教育被认为是促进社会阶层流动的重要途径，而研究型高校中弱势学生往往因其原生家庭的“烙印”，在学业成绩、能力提升（刘志民，高耀，2011）、深造期望等学习结果上常常处于劣势。初代大学生学习结果的劣势在多大程度上受学习过程劣势的影响？目前国内的部分学者深入考察深度学习的群体差异；而在将深度学习作为中间变量，讨论弱势学生与优势学生在深度学习与学习结果关系上差异的学术研究上，仍有较大空间。

因此本文围绕深度学习这一学习过程概念，以比格斯（Biggs, J.B.）的“3-p”模型为理论框架，使用莱尔德（Laird, T. F. N.）和库（Kuh, G. D.）对深度学习高阶、整合、反思三个维度的测量，考察研究型高校本科生深度学习和学习结果的关系。本文进一步从高等教育过程公平视角出发，分析初代大学生和非初代大学生差异，并以深度学习作为中介，研究初代大学生学习结果的劣势。研究旨在揭示研究型高校学生学习过程和学习结果的真实关系，并为提升弱势学生的学习收获提供实证研究上的佐证。

本文以一所研究型大学——F大学为案例，使用2013年5581位本科生的调研数据，采用相关分析、群体差异分析和回归分析方法，得出以下三方面结果：

第一，研究型高校学生深度学习掌握较差，学生相对善于高阶学习，而反思学习和整合学习欠缺；学生技能提升总体欠佳，个人能力提升尚可，人际能力和认知能力提升相对较弱；但学生学业成绩良好，七成学生选择深造。深度学习和学习结果有中等强度的相关关系，深度学习使用频率越高，学生技能提升越好；深度学习三个维度中，整合学习与学生技能提升，尤其是人际和认知能力提升，有较强相关关系。深度学习与深造期望有一定相关关系，但是与学业成绩无关。

第二，与非初代大学生相比，初代大学生在深度学习和学习结果方面均处于劣势。初代大学生整合学习方面劣势尤其严重；认知能力提升、人际能力提升得分更低，并在统计意义上显著；且与非初代大学生相比，初代大学生学业成绩更低，毕业后选择深造的学生比例更低并达到统计意义上的显著。

第三，初代大学生深度学习的劣势是其技能提升处于劣势的重要原因之一，与学生家庭背景和个人禀赋的影响相比，深度学习的影响更大。整合学习对技能提升影响尤其重要。然而在学业成绩和深造期望上，初代大学生家庭背景和个人禀赋

的影响比深度学习的影响更大。

基于以上研究结果，本文对以高绩点作为研究生招生标准的合理性提出质疑；并讨论弱势学生通过深度学习，特别是整合学习应用，突破原生家庭和个人禀赋限制，实现技能提升的可能性。

关键词：深度学习初代大学生学习结果

Abstract: Against the backdrop of rapid development of artificial intelligence that could potentially replace human labors, higher education institutions and employers increasingly value college graduates' general abilities. As an essence of learning process, learning approaches have important effect on college students' learning outcomes and significant implications on higher education quality. Thus, deep approaches to learning (DAL) becomes a hot topic in understanding college student learning. Researches have reached consensus that DAL effects students' ability gains and their educational expectations. Yet they have still debated whether DAL effects grades.

Meanwhile, higher education is broadly regarded as an important path to social mobility, yet reserchers generally find that disadvantaged students gain less in general abilities, have lower educational expectation and underperform academically. However, there lacks of a good understanding of how much these disadvantages might due to students' lower levels of DAL. Only a few researchers examine the group differences in DAL, and explore DAL as a possible mechanism to explain why disadvantaged students have low levels of learning outcomes. Therefore, this study focuses on DAL, guided by Biggs "3-p" model as the conceptual framework and using Laird and Kuh's high-order learning, integrative learning and reflective learning measurements to examine the relationship between DAL and learning outcomes. Furthermore, from the perspective of equality of higher education process, this study examines group differences in DAL and learning outcomes, focusing on first-generation college students. This study aims to explore the connection between students' learning approaches and learning outcomes, and to provide empirical evidences as how to promote general ability gains for disadvantaged students. Using F university as case, based on survey data of 5581 students, correlation analyses, variance analyses and regression analyses have led to the following three main findings:

First, undergraduate students did not use DAL often. Students used high-order learning relatively often but used integrative learning and reflective learning much less. Students general ability gains were limited, especially the cognitive and interpersonal ability gains. Meanwhile, students earned good grades, and over 70% of them planned to study for a graduate degree. There was positive correlation between DAL and learning outcomes. The more often students used DAL, the more they gained in general abilities, particularly cognitive and interpersonal abilities. A weak positive relation existed between DAL and education expectation, yet there was no relationship between DAL and grades.

Second, first-generation college students were in a disadvantaged position in both DAL and learning outcomes. First-generation college students were in a particularly disadvantaged position in integrative learning. They gained much less than

non-first-generation students in cognitive and interpersonal abilities, earned lower grades and had lower educational expectations.

Third, low level of DAL was an important reason first-generation college students gained less in general abilities. The influences of DAL, especially integrative learning, on ability gains were much bigger than that of students' family background and individual characteristics. However, the contrary is true regarding grades and educational aspiration. The influences of students' family background and individual characteristics were much bigger than that of DAL.

In light of these findings, this study questions the validity of using undergraduate grades as a major criteria in graduate school admissions. The findings also imply that disadvantaged graduates need to pay attention to DAL, particularly integrative learning approaches, in order to break the constraints of lower family socioeconomic status.

Keyword: Deep approaches to learning; First-generation college students; Learning outcomes

贫困地区学生高考表现与高等教育选择

刘佳辰 中国人民大学 硕士

摘要: 十九大报告强调扶贫“要与扶志、扶智相结合”，将教育摆在了扶贫工作的突出位置，而教育扶贫的重点是贫困地区学生。在基础教育日益普及的今天，高等教育在学生获取更高教育收益率的过程中发挥着越来越重要的作用。因此，本文将关注贫困地区学生在高考中的考试表现以及在高等教育选择中的决策特征，通过研究贫困地区学生的以上群体性特征，为未来制定更加精准高效的教育扶贫政策提供参考。

本文利用L省2009年高考成绩全样本数据，采用多元线性回归、多元logistic回归、断点回归等计量方法，辅以统计图表的展示，综合研究贫困地区学生的高考表现与高等教育选择，研究内容完整涵盖了学生从参加高考到填报志愿，再到复读决策的全过程，全面分析贫困对学生学业成就与教育选择的影响。

本文研究发现，农村学生在总分和各科分数方面均领先于城市学生，但贫困地区学生在总分、各科分数、一本和二本批次的录取率方面均落后于非贫困地区学生。志愿填报方面，在院校选择上，贫困地区学生的分数利用效率较低，高分低报的现象更加严重；在专业选择上，贫困地区本科学生不愿选择理工科和商科，贫困地区专科学生同样不愿选择商科，但偏好医科和理工科；复读决策方面，在分数略低于二本线的学生中，面对学费更低、声誉更好的二本学校，贫困地区学生的复读率明显低于非贫困地区学生，前者争取更好教育资源的主动性较差。另外，本文还发现学生所在高中的质量对于学生的高考表现、分数使用效率都有显著的正向影响。

本文研究认为，贫困地区学生与非贫困地区学生的差距不止涉及考试表现不佳，

而是一个以低质量高中为诱因，包含了考试表现、志愿填报、信息资源和主动性在内的多层次复杂问题。这启示我们：“农村学生”与“贫困学生”的概念不再高度重合，教育扶贫的政策维度要从“农村—城市”拓展至“贫困—非贫困”，确保扶贫政策的精准高效。二是教育扶贫应以提高贫困地区高中的综合实力为着力点，为贫困地区学生提供包含知识传授、志愿辅导、职业规划、信念鼓舞在内的全方位、多层次的高中教育，以真正实现扶贫与扶“志”、扶“智”的有机融合。

关键词： 高考； 志愿填报； 专业选择； 复读； 教育扶贫

Abstract: The report to the 19th CPC national congress emphasizes that poverty alleviation should be connected with will and wisdom improvement, which endows education with a significant position in poverty alleviation and the key of education poverty alleviation is students. Nowadays, primary education is much more common than ever before and higher education plays a more and more important role in helping students acquire more education rewards. Therefore, we focus on poor area students' exam performance and educational decision in the College Entrance Exam (CEE). Our research on poor area students' features is intended to provide suggestions on more efficient policy of education al poverty alleviation.

The data contains all students' CEE scores and some personal information in a western Chinese province in 2009 and we use several quantitative methods including multiple linear regression, multiple logistic regression and discontinuity regression, and statistic charts and diagrams in order to study the exam performance and education decision of poor area students. What we study on contains the whole CEE procedure from participating in the exam, applying universities, choosing majors and considering whether to retake CEE next year.

We found that compared with urban students, rural students prevail in both total and single subject scores while poor area students lag behind non-poor students in total scores, single subject scores, admission rates of selective universities. In university application strategy, poor area students suffer from lower application efficiency, which means the phenomenon that students with high scores select low quality universities is severe. In major choice, poor area college students dislike STEM and business majors while poor area college students also reject business major but prefer STEM and medical majors. In the decision on retaking CEE, we research on the students near the score requirements of Tier2 university. Despite the Tier2 university has cheaper tuition fee and better quality, poor area students are less likely to retake CEE, which means they are lacking in positive attitude to struggling for better education resource. what's more, we found that the quality of high school has significant impacts on students' exam performance, and application efficiency. We think the gap between poor area students and non-poor students is not just a problem of exam performance but a more complex dilemma involved in high school quality, exam performance, application efficiency, information resource and proactive personality. We can acquire two implications from this. The first is there are differences between rural students and poor students and the education poverty alleviation policy should focus on the poor students in rural area so as to ensure the policy efficiency and social justice. The second is it is the improvement of high

school quality that the education poverty alleviation policy should emphasize and poor area students should be provided with high-quality high school education containing knowledge teaching, instruction on application strategy, career planning and encouragement in order to accomplish the combination of poverty alleviation and will and wisdom improvement.

Keyword: College Entrance Exam; Application Strategy; Major Choice; Retake CEE; Education Poverty Alleviation

职业教育学习评价效用的制度分析——一个链式中介模型的应用实证研究

李鹏 西南大学 博士

摘要: 学习评价是人类重要的制度发现之一,从古代的科举到当代的高考制度,学习评价为教育改革与发展做出了卓越贡献。然而,任何制度都有两面性。学习评价制度在推动并引领教育改革与发展的同时,往往也是教育改革与发展的桎梏。在学习评价的指挥棒作用下,“应试教育”“学业锦标赛”等“社会达尔文主义”的教育乱象层出不穷。如今,教育已从社会边缘走向社会中心,成为了利益交汇的中心,具有高利害性和价值判断属性的学习评价则成为教育改革的问题中心。追问“李约瑟之谜”“韦伯之问”到反思“钱学森之问”发现,学习评价是影响教育体制、人才培养和技术创新的重要制度因素。历史经验与现实困境证明了学习评价制度正经历着严重的“身份危机”——学习评价究竟是教育改革的致命桎梏,还是教育发展的救命稻草?在智能时代、工业 4.0 和教育“中国梦”的时代空间内,聚焦于职业教育领域探究学习评价的“身份危机”。在理论视角上,整合评价效用与制度分析两条线索,综合运用关于学习的评价、促进学习的评价、作为学习的评价等学习评价理论,以及制度分析、制度绩效等制度理论为基础。在研究内容上,以问题为导向,将“学习评价究竟是教育改革的致命桎梏,还是教育发展的救命稻草?”的问题分解为评价效用内涵、生成机理、尺度表征、制度基础、现状水平、影响因素和优化策略等问题。在分析思路,以制度分析的理论框架,依次从制度逻辑、制度变迁、制度运行、制度基础、制度绩效、制度困境、制度创新分析职业教育学习评价效用的具体问题。在研究方法上,采用混合研究的技术路线,综合运用问卷调查法、访谈调查法、课堂观察等数据收集方法,借助 Spss 21.0、Amos 21.0、Mplus 7.0、Excel 2007、Mce 3.2 和 Nvivo 11.0 等研究工具实施职业教育学习评价效用的理论与实践探索。具体研究与探索如下:首先,建构了职业教育学习评价效用的理论框架。一是从“评价促进学习”的教育学机理和“制度生成绩效”经济学机理,整合了“评价实施——结果反馈——教学改进+学习改进——效用生成”的职业教育学习评价效用生成机理。二是从职业教育学习评价制度的经验转向、伦理转向与文化转向批判了评价效用的“工具主义”“空想主义”与“功利主义”取向,明确了职业教育学习评价效用“以改进促发展”的本质,明确了职业教育学习评价效用以人为本、公平与发展的价值理

性。其次，评判了职业教育学习评价效用的制度基础。一是运用开放式访谈、文献分析等方法自编《职业教育学习元评价调查问卷》，基于“群组层次分析法”确定系数权重和主体评分权重，合成职业教育学习元评价结果等级。二是采用分层随机抽样法则在全国实施元评价调查，发放专家问卷40份，有效回收32份；发放学生问卷400份，有效回收337份；发放企业问卷150份，有效回收103份。结合专家、学生和企业代表评分，对职业教育学习评价的实用性、可行性、合理性与准确性进行了评价。三是运用沃斯特根指数测算职业教育学习评价的“起点——过程——结果”公平指数，判断职业教育学习评价制度实施的公平状况。然后，测度了职业教育学习评价效用的实然状况。一是结合职业教育学习评价效用的过程分析框架建构了学习评价、教学改进、学习改进和KSA发展的链式中介模型。二是综合运用文献法、访谈法，编制《职业教育学习评价效用测度问卷》，并借助项目分析、探索性因素分析和验证性因子分析对问卷进行了测量学性能分析和修订。三是采用分层随机抽样在全国实施问卷调查，发放问卷2132份，回收有效问卷1867份。四是通过发展指数、多维比较与中介效应验证分析了职业教育学习评教效用的实然状况。再次，探索了职业教育学习评价效用的影响因素。在量化分析的基础上，结合量化分析的问题和假设，采用质性分析的方法进行问题的原因分析。先后对10名职业教育教师和18名学生进行现场访谈或电话访谈，然后借助NVivo 11.0软件对访谈结果进行自由编码、主轴编码和选择性编码。在完成编码的基础上，结合制度绩效影响因素和学习评价影响因素的相关理论，从宏观、中观、微观三个层次分析了影响职业教育学习评价效用的文化、理性与心理的多重因素。最后，提出了职业教育学习评价效用提升的对策。结合元评价发现的制度运行问题、现状测度发现的效用困境以及质性访谈的多重原因分析，从提高职业教育学习评价效用的视角出发，提出了提高职业教育学习评价效用的对策建议：一是变革职业教育学习评价制度，二是推进学习评价后的教师教学改进，三是引导学习评价后的学生学习改进。通过理论分析、实证分析与假设验证，得到关于职业教育学习评价效用研究的主要发现：第一，职业教育学习评价效用的生成机制是“评价实施——结果反馈——教学改进+学习改进——效用生成”的过程系统。职业教育学习评价效用是评价制度运行的结果，在本质上是“以评价促进改进，以评价促进发展”，在结果上体现为学生的KSA发展。在生成机理上，职业教育学习评价通过制度运行生成客观的评价结果；基于结果分析与反馈，促进教师教学改进和学生学习改进；通过教学改进与学习改进的共同作用，最终促进职业教育学生的KSA发展。第二，职业教育学习评价效用观是发展变化的，但始终坚持着人本、公平与发展的价值尺度。职业教育学习评价效用观随着评价历史的发展而不断变化，职业教育学习评价是为了培养“职业人”“道德人”和“匠心人”，包含了国家“立德树人”的政治理想、学校与教师“教书育人”的本质工作、学生与家长“追求发展”的投资行为、企业“人才选拔”的参考依据等多重治理逻辑。在实践中，职业教育学习评价制度始终要坚持人本、公平与发展的价值理性。第三，职业教育学习评价制度运行情况较好，但也存在着实用主义、技术主义和形式主义的问题。职业教育学习评价制度的沃斯特根指数为 $1.5793 > 1$ ，整体运行较好。但也存在着三个方面问题，一是职业教育学习评价制度起点公平得到重视，但是制度逻辑倾向实用，起点公平指数为 $1.6151 > 1$ ；二是过程公平得到彰显，但是具体流程依赖技术，过程公平指数为 $1.6320 > 1$ ；三是结果公平得到认同，但是实际行动流于形式，结果公平指数为 $1.7008 > 1$ 。第四，职业教育学习评价效用的整体发展指数尚可，但绝对水平偏低。职业教育学习评

价效用的综合发展指数为 $0.5140 < 1$ ，从 0.5140 到 1 之间还有巨大的改进和发展空间。而且，职业教育学习评价对 KSA 发展的路径作用系数仅为 $\beta=0.46$ ，直接效应值为 0.464 ，总效应值 0.653 。第五，职业教育学习评价效用的局部差异特征显著，但总体关系相对稳定。方差分析与多维均值比较发现不同区域、不同级别和不同类型职业院校中，学习评价效用的差异性特征比较明显。但学习评价、教学改进、学习改进与 KSA 发展之间的关系相对稳定，链式中介模型检验的拟合指数良好，各变量值之间的数量关系稳定。第六，职业教育学习评价效用的中介效用明显，但内部结构需要完善。教学改进在学习评价与 KSA 发展之间的中介效应为 0.106 ，效果量为 16.23% ；学习改进的中介效应为 0.042 ，效果量为 6.43% ；教学改进+学习改进的中介效应为 0.041 ，效果量为 6.28% 。但是教学改进的中介效应(作用系数 0.106 ，效果量 16.23%) 优于学习改进的中介效应(作用系数 0.042 ，效果量 6.43%)。第七，从制度运行的条件来看，影响职业教育学习评价效用的因素主要包括文化、理性和心理三个方面。宏观层面的文化因素和制度环境，中观层面上制度结构、制度要素与人的理性，微观层面上制度内部人与人之间的交往心理和博弈等都会影响职业教育学习评价效用。

关键词：职业教育；学习评价；评价制度；评价效用；中介效应；制度分析

Abstract: Learning evaluation is one of the important institutional discoveries of human beings. From the *imperial examination* to the *college entrance examination*, learning evaluation has made outstanding contributions to the reform and development of education. However, each institution has two sides. The learning evaluation institution is also the shackle of education reform when it promoting education develop. Under the baton and guidance of learning evaluation institution, the *examination-oriented education*, *academic championship* and *social Darwinism* emerges endlessly in modern education. Therefore, the difficulties of education equity and development seem to be caused by learning evaluation institution. Nowadays, education has moved from the social fringe to the social center, and with high interest and value judgment attribute, learning evaluation institution also become the problem center of education reform. From “*Needham Puzzle*” to “*the question of Qian Xuesen*”, the evaluation of learning is an important institutional factor affecting the educational system, as well as an important breakthrough in the education reform. The historical experience and the realistic predicament have proved that the learning evaluation institution is experiencing a serious “identity” crisis. So, the learning evaluation is the fatal shackle of educational reform, or the lifeline of educational development? In the era of intellectual age, industry 4.0, this paper focuses on the “identity” crisis in vocational education. Based on the theories of *assessment of learning*, *assessment for learning*, and *assessment as learning*, also the institutional performance and institutional analysis theory, decompose the performance of Learning Evaluation in Vocational Education. And divide the main research problem into several sub-problems, performance connotation, generation mechanism, realistic foundation, current situation level, influence factors and optimization strategy. Based on the theoretical framework of institutional analysis, this paper analyzes the performance of learning evaluation in vocational education from institutional change, institutional logic, institutional structure, institutional operation, institutional

performance, institutional dilemma and institutional innovation. Integrating the two clues of performance problem and institution analysis, adopting the technical route of mixed research, synthetically using questionnaire survey, interview investigation, classroom observation and other data collection methods, with the help of Spss 21.0, Amos 21.0, Mplus 7.0, Excel 2007, Mce 3.2 and Nvivo 11.0 and other research tools to carry out the exploration of the performance of Learning Evaluation in Vocational Education. The specific research and exploration are as follows: Firstly, this paper constructed the theoretical analysis framework of the performance of Learning Evaluation in Vocational Education. By analysis the turning of experience, ethics and culture of the learning evaluation institution of vocational education, this paper criticizes the orientation of “*instrumentalism*” “*utopism*” and “*utilitarianism*”, and then clarified the consistency of “*teaching-learning-evaluation*” in vocational education and the essence of “improving and promoting development”. Then, this paper constructed the mechanism of evaluation performance generation from the process of institution operation “evaluation implementation-result feedback-teaching improvement + learning improvement-performance generation”. In addition, this paper found the value basis and institution design logic of the performance of Learning Evaluation in Vocational Education, the value demand and governance logic of stakeholders, the value rationality of humanism, fairness and development. Secondly, this paper evaluates the practical basis of the performance of Learning Evaluation in Vocational Education. Using the open interview, literature analysis and other methods, developed the *learning meta-evaluation questionnaire in vocational education*. Using Mufti-AHP method to count the weight of coefficient and subject scoring, and then composing the grade of learning meta-evaluation result in vocational education. Using the stratified random sampling law to carry out meta-evaluation nationwide survey, distributed 40 expert questionnaires and collected 32 copies valid questionnaires, distributed 400 students questionnaires and collected 337 copies valid questionnaires, distributed 150 copies of enterprise questionnaires and collected 103 copies valid questionnaires. Combining with the scores of experts, students and representatives of enterprises, meta-evaluated the *practicability, feasibility, rationality and accuracy* of the learning evaluation in vocational education. By using *Verstegen index*, this paper calculate the “starting-process-result” fairness index, and judge the fair situation of the learning evaluation institution in vocational education. Thirdly, this paper measured the actual status of the performance of learning evaluation in vocational education. Using processing analyze to construct the framework of the performance of learning evaluation in vocational education, construct the chain intermediary model of learning evaluation, teaching improvement, learning improvement and KSA development. Comprehensive use of literature method, interview method, compiled the *performance measure questionnaire of learning evaluation in vocational education*, and revised the questionnaire byte help of item analysis, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Using the stratified random sampling law to carry out questionnaire survey in the whole country, distributed 2132 questionnaires and collected 1867 copies valid questionnaires. Through the development index, the multidimensional comparison

and the intermediary effect verification, the actual performance of the learning evaluation in the vocational education was been analyzed. Fourthly, this paper explored the influencing factors on the performance of learning evaluation in vocational education. Combine the theory analysis and the qualitative data methods to explore the causes of the performance of learning evaluation in vocational education. Based on the interviews of 10 vocational education teachers and 18 students, and combine the data coding and theoretical experience, this paper analyzes the influence of factors on the performance of learning evaluation in vocational education from macro, meso and micro level. Based on the causes of the institutional dilemma, this paper explored the relationship of the performance of learning evaluation in vocational education and culture factors, rationality factors and psychology factors. Finally, the paper puts forward the countermeasures of improving the performance of learning evaluation in vocational education. Combined with the institution operation problems found by meta-evaluation, and the performance dilemma of current situation measurement and the multiple causes of qualitative interviews, this paper puts forward the countermeasures and suggestions to improve the performance of learning evaluation in vocational education: reforming the learning evaluation institution in vocational education, promoting the teachers' teaching improvement, and guiding the students' learning improvement after learning evaluation. The main research findings of the study on the performance of learning evaluation in vocational education are obtained through theoretical analysis, empirical analysis and hypothesis verification: Finding one, the performance of learning evaluation in vocational education includes various institutional logics while the most fundamental measure of value is humanism-oriented, fair and development. The learning evaluation in vocational education is aimed at cultivating "professional people", "moral people" and "craftsmen", including the political ideal of the state of "building morality and cultivating people", the essential work of schools and teachers "educating and educating people", the investment behavior of students and parents to pursuit "development", the reference basis of enterprise "talent selection". But the fundamental value measure of vocational education and learning is humanism, equity and development. Finding two, the generating mechanism of performance of learning evaluation in vocational education is the process system of "evaluation implementation-result feedback-teaching improvement + learning improvement-performance generation". In terms of generating mechanism, learning evaluation of vocational education generates objective evaluation results through institution operation (direct performance), based on result analysis and feedback, to promote teachers' teaching improvement and students' learning improvement (intermediary performance and finally promote the vocational education students' KSA development (ultimate performance)). Finding three, the evaluation institution of vocational education and learning works well, but there are some problems of pragmatism, technical's and formalism. The *Verstegen index* of learning evaluation institution in vocational education is $1.5793 > 1$, so the whole operation is OK. However, the starting point equity of learning evaluation institution in vocational education tends to be pract, the starting point equity index is $1.6151 > 1$. The process

fairness is obvious while the concrete process depends on the technology, and the process equity index is $1.6320 > 1$. The result fairness index is $1.7008 > 1$, while the actual action is mere formality. Finding four, the development index of the performance of learning evaluation in vocational education is fair, while the absolute level is low. The study found that the comprehensive development index of performance of learning evaluation in vocational education is $0.5140 < 1$, and there is great room for improvement from 0.5140 to 1. And the role of Vocational Education Learning Evaluation on the path of KSA development is only 0.46 while the directly effect value is 0.464 and the total effect value is 0.653. Finding five, the characteristic of the performance of learning evaluation in vocational education is difference, while the whole level is relatively stable. The multi-dimensional mean comparison found that the performance of learning evaluation in vocational education have difference characteristics are obvious in statistically. However, the relationship among learning evaluation, teaching improvement, learning improvement and KSA development is relatively stable while the fitting index of chain intermediary model test is good, and the quantitative relationship between variables is stable. Finding six, the Intermediary effect is obvious in the function mechanism of the performance of learning evaluation in vocational education, while the internal structure needs to be perfected. It is found that the mediating effect between teaching improvement and KSA development is 0.106 while the effect is 16.23%; the mediating effect of learning improvement is 0.042 while the effect is 6.43%; the mediating effect of teaching improvement and learning improvement is 0.041 while the effect is 6.28%. However, the mediating effect of teaching improvement (the effect coefficient 0.106, the effect quantity 16.23%) is better than the mediating effect of learning improvement (0.042, 6.43%). Finding seven, from the perspective of institutional analysis, there are many factors that affect the performance of learning evaluation in vocational education. At the meso level, the cultural factors and the institutional environment affect the performance. Also, at the macro level, the institutional structure, the rationality of the institutional elements and the rationality of the people in the system. And at the micro level, the interaction psychology and the game among the people within the institution will all affect the performance of learning evaluation in vocational education.

Keyword: Vocational education; Learning evaluation; Evaluation institution; Evaluation performance; Intermediary effect; Institutional analysis

技术促进小学数学深度学习的实证研究

胡航 西南大学 博士

摘要: 当前,国家对教育信息化的研究和实践给予了大量人力、物力和财力,然其是否真的促进了学习者发展,学校教育并未完全给出积极的反馈。小学数学作为学校教育的核心构成,其学习品质受到研究者和实践者的广泛关注。研究围绕“技

术促进深度学习”这一核心命题,以小学数学为载体,深入学科内部,探索技术促进学科教学的有效范式。据此,确定了研究问题:一是小学数学深度学习品质如何表征?二是技术如何促进小学数学深度学习?三是技术促进小学数学深度学习的机理是什么?研究运用文献分析、准实验研究、实验研究、问卷与访谈等方法,经历了四个研究阶段:第 1 部分文献分析。确立了研究问题、研究方法、研究思路、研究框架等内容(绪论),厘清了 CTCL 研究范式、深度学习和小学数学教学的研究现状及其关系(第一章)。第 2 部分诊断性研究。运用已有量表,聚焦小学生数学学习策略和认知结构,开发了小学数学可视化诊断系统,对四城市小学生数学学习品质现状进行了评测,为实证研究提供了评估工具和方法(第二章)。第 3 部分实证研究。运用准实验和实验研究,在不同水平学校和不同年级中先后进行了 3 个课堂实验,每个实验均每天 1 学时,持续 1 个月。实验对学业成绩、学习策略、认知结构、眼动、ERP 脑电、学习者注意力与学习者互赖关系等数据进行了分析与比较,具体如下:一是多变量探索性实验,确立了技术促进小学数学深度学习的两个主要影响因素“学习内容”与“学习方式”(第三章);二是针对“学习内容”的验证性实验,其变量为“深度学习内容构成及其数字化资源表征”;三是针对“学习方式”的验证性实验,其变量为“个性化-合作”学习(第四章)。第 4 部分范式构建与理论总结。通过实证研究,归纳了技术促进小学数学深度学习的操作范式(第五章),包括针对“学习内容”的“4S”学习内容构成、课程重构框架、S-DIP 资源表征态、CRF 资源开发模式;针对“学习方式”的 PCL 设计原型与 PCL 模型等;解析了深度学习发生过程与设计模型 S-ACIG,析出技术促进小学数学深度学习的发生机理,明确了深度学习中的技术使能,形成了课堂深度学习的实践框架(第六章),并总结了整个研究的结论、成果、不足与下一步研究计划(结论与展望)。针对三个研究问题,研究得出了以下结论:一是认知(学业成绩等)、生理(眼动等)和情感(学习者互赖关系等)三个层面共同表征深度学习;二是“4S”学习内容构成及其资源表征促进了知识的多样态呈现并与深度学习过程契合;“个性化-合作”学习方式促进学习者基于个性的社会化过程发生;三是“分布式交互中数形转化”是技术促进小学数学深度学习的发生机理。研究中实验组在认知、生理和情感三个层面都获得了较显著的发展;操作范式具有较强的可推广性,丰富了技术促进学科学习的理论与实践。

关键词: 小学数学;深度学习;学习内容构成;数字化资源表征;“个性化-合作”学习

Abstract: At present, the state has given a large amount of manpower, material and financial resources to the research and practice of educational informatization. However, whether it has really promoted the development of learners, school education has not fully given positive feedback. The mathematics in primary school is one of the core components in school education, and its learning quality has been paid close attention to by researchers and practitioners. The study focuses on the core content of "technology promoting deeper learning", and tries to solve three problems. Firstly, How to characterize the quality of deeper learning on mathematics in primary school? Secondly, how does technology promote deeper learning on Math. in primary school? Thirdly, what is the mechanism for technology to promote deeper learning on Math. in primary school? The research mainly uses 5 research methods: literature research, quasi-experimental research, experimental research, questionnaire and interview. For the above three research

questions, the research has gone through four stages:

Part 1 Literature Review. In the stage, the study establishes research questions, research methods, research paths, research framework and so on (introduction), clarifies the status of research at home and abroad about CTCL researcher-centered design and deeper learning and the logical relationship between paradigm, between the three (Chapter 1).

Part 2 Diagnostic Study. In the stage, the study uses the existing scales, focuses on the learning strategies and cognitive structure on Math. in primary, and develops the visual system and diagnosis the deeper learning quality on pupils, provides assessment tools and methods for the empirical research. (Chapter 2).

Part 3 Empirical Research. In the stage, Using quasi experimental and experimental research, 3 classroom experiments were conducted in different grades in different levels schools. Each experiment lasted 1 hour every day for 1 month. the study analyzes and compares the data about academic achievement, learning strategies, cognitive structure, eye-movement, ERP EEG, learner attention and the interdependence relationship between learners. As follows: through multivariate exploratory experiment, the study establishes the two main effect variables—"learning content" and "learning style" (Chapter 3); which has a univariate verification experiment on the "learning content", which is embodied in the "learning content constitutes and its digital resource presentation"; which embodied in the "personalized - cooperative" learning by the univariate verification experiment on the "learning style" (Chapter 4).

Part 4 Paradigm construction and theoretical summary. In the stage, according to the empirical research, the study sums up and summarizes the operational paradigms of deeper learning in primary school on mathematics. As are included: the "4S" learning content structure, UDA curriculum reconstruction framework, S-DIP resource representation, and CRF resource development model for "learning content"; the PCL design prototype and PCL model for "learning style"(Chapter 5). Then this paper analyzes the deeper learning process model S-ACIG, summarizes the mechanism of the technology to promote the deeper learning in class (mathematics), and constructs the practical framework of the classroom learning (Chapter 6), and sums up the conclusions of the study, achievements, disadvantages and the next step plan (conclusions and outlook). Aiming at the above three problems, The study draws the following conclusions: in the first place ,cognitive, physical and learning psychology are the joint characterization of deeper learning. Next, through the following four aspects to promote deeper learning on mathematics in primary school: "4S" learning content and "S-DIP" resources presentations promote the diversification representation of knowledge; "personalized-cooperative" learning improve socialization based on personalization. In the end, "the transformation between data and shapes in distributed interaction" is the mechanism of technology to promote the deeper learning in class. The empirical research occurred in different grades in different levels schools, and the experimental group have gained significant

development in three aspects of cognitive, physical and emotion. So operation paradigm can be popularized, and the research achievements enrich the theory and practice of technology promoting deeper learning.

Keyword: the Math. in primary school; deeper learning; learning content constitution; Digital resource representation; “personalization – cooperative” learning

基于行为过程表现测量合作问题解决能力的研究

袁建林 北京师范大学 博士

摘要: 测量是现代科学的基础,教育测量同样是教育科学的基础。但在教育的人文取向和科学取向争论中,教育在科学化的道路上步履维艰、发展缓慢。制约教育科学化进程的重要因素之一是教育测量发展的滞后,对于教育中的抽象概念和宏大命题,缺乏有效手段和精确模型将其量化,因而难以基于科学的范式对其进行研究。核心素养是近年来教育研究的热点问题,对于这类高度抽象、结构复杂的高阶技能,如何科学而有效的对其进行测量,是教育测量领域面临的重要挑战。合作问题解决能力是一种重要的核心素养,同样面临着测量方法层面的难题,虽然PISA和ATC21S分别采用“人机交互”和“人人交互”的方式开展了大规模测评研究,但都存在一定局限和不足。在教育测评发展层面,过程性评价是一种全新的评价理念,因其关注学习过程、关注任务完成过程,能获取真实的过程表现和详细的过程数据,是解决复杂技能测量问题的可行性途径,但因缺少系统的方法过程和成熟的操作模式,过程性评价更多地表现为一种评价理念和方法导向,较少落实到实际测验中,“过程性测量”得不到体现。

本研究置身于测量方法革新和复杂技能的测量需要这个双重背景,以过程性测评理念为指导,探索基于行为过程表现测量学生的合作问题解决能力。研究主要包含两个层面的目标,一方面,厘清基于行为过程表现进行测量的基本内涵和方法过程;另一方面,基于该“过程性测量”的模式对合作问题解决能力进行科学、有效的测量。本研究综合使用了文献研究法、行动研究法、访谈法、问卷调查法,以及项目反应理论、验证性因子分析、结构方程模型等数据处理技术;首先界定了“基于行为过程表现进行测量”的基本内涵,认为其是过程性评价理念下发展起来的一种新测量范式,它的关键特征包括:强调复杂技能的测量,需要构建动态性、交互性的测验任务,能够获取复杂、详尽的测验过程表现,需要从行为过程表现中抽取测量证据进行基于证据的推理;它需要“证据中心的设计”理论指导测验设计与开发,围绕“评价是基于证据进行推理”的核心思想开展各项工作;在此基础之上主要完成三方面研究工作。主要研究过程和成果如下:

第一,设计并开发合作问题解决能力测验任务。明确了任务设计分为任务原型设计和测验表现形式设计两部分。利用“信息不对称”和“达成共同目标”的二元机制诱发学生合作性地完成共同任务。提出“问题情境的结构机制”,通过界定任务原型中“关键事件”对问题情境中的活动及问题情境的状态进行抽象,“关键事件”在时间维度上的连续构成测验的过程表现,即行为过程流。本研究共设计和开发了5个任务,认知任务分析的结果表明,所有任务能有效地诱发学生参与合作问题

解决过程。所有任务采用统一的界面呈现形式和单线型执行流程，测验系统经过多轮的修改，有较好地协同机制、运行稳定，能应用于实际大规模测验。

第二，基于行为过程表现抽取测量证据。本研究明确复杂表现性测验中要区分测验数据与测量证据的关系，需要先从复杂测验表现中抽取测量证据。在合作问题解决能力测验中，对构念元素进行操作性界定是测量证据建构的逻辑起点，以此定义构念元素映射于具体任务中的具体行为模式或行为序列，从而获得具体任务中学生能力的观测项目。本研究针对“两人交互”的测验模式，提出“个人项目”和“小组项目”两种类型共存的项目设计思路，完成四个正式测验任务的项目设计。邀请了五位专家对所有项目进行评审，评审结果表明整体测验项目设计较为合理，并依据专家反馈意见完成了修改工作。根据项目设计，编写了人工算法程序实现对行为过程流的自动化评分，最后包含个人层面项目44个，小组层面项目8个。对项目的评分一致性研究表明，评分者之间一致性程度较高，评分者与人工算法程序之间的一致性也较高，测验项目设计合理。

第三，基于新测量模型分析测验项目质量和测验结果。本研究采用项目内多维度Rasch模型对测验项目进行了两轮分析，最终测验项目包含个人层面项目36个，小组层面项目8个。所有项目的MNSQ指标表现较好，项目区分度处于可接受范围，测验项目的难度均值为-0.107，测验的总体难度稍显容易，所有个体层面的项目与其对应的任务目标项目之间显著正相关。基于两维度Rasch模型和四维度Rasch模型估计学生能力，结果表明合作小组两位学生的合作问题解决能力相关为0.561，合作小组成员存在较高相互依赖性，学生的合作能力与问题解决能力相关系数为0.468，合作小组两位成员的合作能力与问题解决能力存在较复杂的相关关系。基于结构方程模型建立同伴互评的合作问题解决过程表现对合作问题解决测验结果的预测关系，双因子预测模型的预测效度分别为0.209和0.314，说明本测验具有较好的效度。

本研究表明，基于行为过程表现测量复杂技能是一种可行的途径，基于行为过程表现测量合作问题解决能力具有较高的测量学质量。本研究的创新之处在于：初步较系统地提出“基于行为过程表现进行测量”的概念与方法过程，提出了复杂测验任务设计的基本方法，阐释了复杂测验表现中抽取测量证据的原理，基于新的思路设计合作问题解决能力测验项目和建构新测量模型。同时，需要不断完善“过程性测量”的理论与方法体系，需要拓展“过程性测量”的应用面，需要充分利用机器学习与人工智能中的新方法丰富测量证据的抽取，更深层次探讨合作分组、测验项目设计和测量模型建构等问题，从而使得本研究更加完善和丰满。

关键词：过程性测量；合作问题解决能力；行为过程表现；证据中心的设计

Abstract: Measurement is the basis of modern science, and educational measurement is also the foundation of educational science. In the debate on the humanistic orientation and scientific orientation of education, education is struggling and developing slowly on the scientization path. One of the important factors that restrict the process of scientizing education is the lag of the development of education measurement. The lack of effective means and precise models that might be used for quantifying Abstract concepts and grand propositions in education makes it difficult to study them based on the scientific paradigm. Key competencies is a hot issue in educational research in recent years. For these high-level skills that are highly Abstracted and complexly structured, how to scientifically and effectively measure

them is an big challenge in educational measurement field. Collaborative problem solving is an important skill in all key competencies, which also faces the problem of how to quantify it. Although PISA and ATC21S have carried out large scale program for assessing collaborative problem solving separately based on human-to-agent approach and human-to-human approach, both of them exist some limitation and shortness. In the aspect of development of educational assessment, process assessment is a new concept of evaluation. Because it focuses on the learning process, the process of completing task, and it can obtain actual process performance and detailed process data, process assessment is a feasible approach to solve the problem of measuring complex skills. However, because of the lack of systematic methodologies and reliable model of executing process test, the process assessment is showed as a kind of assessment concept in more cases, less implementation to the actual test, and the “process measurement” is not reflected.

This research is based on the dual background of measuring method innovation and requirement for complex skills measurement. Guided by the concept of process assessment, we explore the methodologies for measuring students' collaborative problem solving based on behavioral process performance. The study mainly includes two objectives. On the one hand, it is needed to clarify the basic implication and methods of measuring based on the behavioral process performance; on the other hand, based on the "process measurement" mode, we conducts a scientific and effective practice for measuring collaborative problem solving. This study integrates usage of the research methods such as literature review, action research, interview, questionnaire survey, and data analysis technique such as item response theory, confirmatory factor analysis and structural equation modeling. At the beginning, the study clarify the basic implication of measurement based on behavioral process performance, which is a new measurement paradigm developed under the concept of process assessment. Its key features include emphasizing the measurement of complex skills, requiring dynamic and interactive testing tasks, obtaining complex and detailed test process performance, and requiring extracting evidence from behavioral process performance then conducting evidence-based reasoning. It requires Evidence-centered Design theory to instruct test design and development, and carries out various test work around the core idea of “assessment is a special case of evidentiary reasoning”. Based on the work of concept clarification, we conduct tree major researches. The research process and results are as follows:

First, task design and development for measuring collaborative problem solving. We divide the task design into two parts: the task prototype design and the test performance design. The dual mechanism of “information asymmetry” and “to achieve common goals” is used to induce students to complete tasks collaboratively. The study put forward the structure mechanism of problem situation. Based on this mechanism, we can Abstract problem situation activities and problem situation state through defining “critical event” in the task prototype. The test process performance is continuously constituted by “critical event” in the time dimension, which is also called behavioral process flow. Five tasks are designed and developed in this study. The results of cognitive task analysis show that all tasks could effectively induce

students to engage the process of collaborative problem solving. All tasks have unified interface and single line execution process. After several rounds of modification, the test system has better coordination mechanism and stable operation, and can be applied to large-scale practical tests.

Second, extracting measurement evidence from the behavior process performance. In this study, we claim that the relationship between the test data and the measurement evidence should be distinguished in the complex performance test, and the measurement evidence should be extracted from the performance of the complex test. In collaborative problem solving test, the operational definition of construct element is the logical starting point for constructing measurement evidence, then identify behavior patterns in specific tasks that mapping construct element, so as to obtain observational indicators of students' ability in specific task. Toward the test mode of "two person interaction", this study define two type items: "personal item" and "group item", and complete item design of the four official test tasks. Five experts are invited to review all the items. The results show that the overall item design is more reasonable, and the revision work is completed according to the experts' feedback. According to the item design, we compile the artificial algorithm program to achieve the automatic scoring of behavioral process stream, and finally include 44 personal level items and 8 group level items. The study of scoring consistency shows that the consistency between raters is high, and the consistency between rater and artificial algorithm program is also high. The design of the test items is reasonable.

Third, analyzing item quality and test results based on new measurement model. In this study, we conducted two rounds of analysis on the items using the within-item multidimensional Rasch model, and the final items consists of 36 individual level items and 8 group level items. The MNSQ of all item have a good performance, and the discrimination of all item is acceptable. The average difficulty of all item is -0.107, that means the overall difficulty of the test is slightly easy. There are significant positive correlations between all the individual level items and the corresponding task target items. Respectively based on two dimensional Rasch model and four dimensional Rasch model to estimate the student ability, the results show that in collaborative group, the correlation coefficient of the two students' collaborative problem solving ability is 0.561, that means the team members have a high dependencies to each other. The correlation coefficient of students' collaboration ability and problem solving ability is 0.468, and the correlation of collaboration and problem-solving of the two members in collaborative group is more complex. Based on the structural equation model, the process performance of collaborative problem solving that conducted by peer review is used to predict the results of collaborative problem solving test, and predictive validity of the double factor prediction model are 0.209 and 0.314, that means the test has good validity.

This study shows that it is a feasible way to measure complex skills based on behavioral process performance. Measuring collaborative problem solving based on behavioral process performance shows good testing quality. The innovations of this research are preliminarily clarifying the concept and systematical method of

measurement based on behavioral process, clarifying how to design complex tasks, and illustrating the principles of extracting the measurement evidence from the complex test performance, and based on the new ideas, the test item and new measurement model are designed for measuring collaborative problem solving. At the same time, it is need to continuously improve the theory and method system of “process measurement”, to expand the application of “process measurement”, and to make full use of the new method of machine learning and artificial intelligence so as to enrich the method of evidence extraction, to deeply explore problems such as group composition, item design and measurement model design, and all these work will make the current research more perfect.

Keyword: Process Measurement; Collaborative Problem Solving; Behavioral Process Performance; Evidence-centered Design

文化生态视野下藏族中学理科教育问题研究

党宝宝 西北师范大学 博士

摘要: 少数民族理科教育质量关系到民族地区理工科人才的培养和区域经济发展的整体实力,对于优化民族地区的人力资源,均衡民族教育文化生态环境具有重要的意义。目前,少数民族理科教育问题集中显现在基础教育领域。研究以甘南藏族自治州的藏族中学为研究对象,采用定性(参与式观察法、访谈法、内容分析法)与定量(问卷调查法)相结合的混合式研究方法,依据文化生态理论,微观层面上,重点从藏族中学生的学习特点、语言学习环境、理科教师配置、理科教学过程、理科教材建设等方面进行剖析;宏观层面上,主要分析了藏区理科教育发展的社会文化困境。通过对藏族中学生理科学习现状的分析,研究发现,甘南州藏族中学理科教育发展相对缓慢,学生理科学业水平普遍低下。无论是初中生还是高中生,都存在理科课程学习困难的问题。对藏族中学生理科学习特点的分析,结果表明,藏族中学生普遍持有较低的理科学习兴趣和理科学习自我效能感水平,并且,在性别、聚居区域和年级上均存在显著的差异效应。通过对藏族中学生理科学习的作用机制分析,研究发现,藏族中学生理科学习的文化心理尺度在理科学习兴趣和自我效能感之间存在链式中介作用,表明理科学习中的外围文化因素会影响他们对理科课程以及理科学业效能的态度。藏族中学普遍实施“一类藏汉双语教育模式”,理科课程均使用藏语授课,藏汉双语教育形式对理科教育产生基础性影响。研究发现,藏族中学生对藏语理科授课普遍持有积极的态度,具有较高的诉求性。藏汉双语学习态度和理科学习倾向性之间存在显著的相关性,且存在显著的性别和年级差异效应,回归分析发现,藏汉双语学习态度会影响他们对理科知识和课程学习的倾向性。研究也发现,藏族中学生偏爱理科藏语学习,有其深刻的社会缘由,例如,藏语理科学习模式和机制的养成,对汉语理科知识的认知困难,对母语的认同偏好和先验知识的影响等。通过对藏族理科教师、藏语理科教材和藏语理科教学过程三个核心元素的分析,研究发现,藏族中学的教师资源配置存在不合理现象,理科教师配置出现结构性失衡,跨专业任课现象普遍。藏族理科教师的理科教学基础能力和延展能力有待提高,理科教

师的培训机制处于被动局面，信息化教学素养欠缺。同时，学期内实验课程开设次数较少，实验课程以课堂演示为主。理科教材的编译和结构安排上忽视了藏族中学生的认知发展特点，案例与素材缺少藏文化特色，课程标准缺乏弹性，理科教材难度偏大。研究认为，藏族中学理科教育发展也受到社会文化生态因素的困扰。在区域社会文化氛围生态层面上，藏区社会理科学习文化气氛淡薄，理科知识远离藏区生活实际，理科教育的功能无法与藏区社会的实际需求相适应，同时，寺院教育对现代学校教育产生强烈冲击；在家庭教育文化生态层面上，藏族中学生家庭教育严重缺失，传统的“重文轻理”家庭教育观和性别教育观影响藏族学生对理科教育的态度；在个体文化心理生态层面上，藏族中学生独特的民族品性会影响他们理科学习的倾向性；在学校制度生态层面上，民族院校文理科专业设置的结构性失衡诱发藏族中学文理分科的偏态化倾向，独立封闭的民族教育体系限制了藏族中学生的理科教育途径，理科教师职业发展的恶性环形路径降低了教师教学水平，积极的高考优惠政策对学生理科学习产生的实际效用较小。因此，甘南藏区的理科教育应实施适宜的藏汉双语教育模式，即混合式藏汉双语教育模式，着力提升藏族理科教师的专业素养和教学能力，打造有效的理科课堂教学模式，编译符合藏族学生认知发展特点的理科教材，同时，积极推进藏族理科教育的文化生态环境建设。在此基础上，结合研究的局限性，提出了将来研究的趋向。

关键词：藏族，理科，中学生，文化生态

Abstract: The quality of science education for ethnic minorities is related to the cultivation of science and engineering talents in the minority areas and the overall strength of regional economic development. It is of great significance for optimizing human resources in ethnic areas and balancing the cultural and ecological environment of ethnic education. At present, the issue of education for ethnic minorities is concentrated in the field of basic education. The study used Tibetan middle schools in Gannan Tibetan Autonomous Prefecture as the research object, using a hybrid research method combining qualitative (participatory observation, interviewing, content analysis) and quantitative (questionnaire survey) methods, based on cultural ecology theory, on the micro level, focusing on the study characteristics of Tibetan middle school students, the language learning environment, the allocation of science teachers, the teaching process of science, and the construction of science textbooks; the macro-level analysis mainly focuses on the socio-cultural dilemmas of the development of science education in Tibetan areas. Through the analysis of the current state of physical education in Tibetan middle schools, the study found that the development of science education in Tibetan middle schools in Gannan Prefecture is relatively slow, and the academic level of students in science is generally low. Both junior high school students and high school students have problems in learning science courses. The analysis of the characteristics of the physiology of Tibetan middle school students shows that the physiology department of Tibetan middle school generally holds lower levels of science interest and science self-efficacy, and there are significant differences in gender, inhabited areas and grades. Through the analysis of the mechanism of physical education in Tibetan middle schools, the study found that there is a chain-mediated effect of the cultural and psychological scales of physical learning in Tibetan middle schools between science interest and self-efficacy, indicating that the peripheral cultural factors in science study will affect their Science

curriculum and attitudes toward the efficacy of science. Tibetan middle schools generally implement “One Type of Tibetan-Chinese Bilingual Education Model”, science courses use Tibetan language teaching, and Tibetan-Chinese bilingual education forms have a fundamental impact on science education. The study found that Tibetan middle school students generally hold a positive attitude towards Tibetan science teaching and have high appeal. There is a significant correlation between Tibetan and Chinese bilingual learning attitudes and science learning orientations, and there are significant gender and grade differences. Regression analysis finds that Tibetan-Chinese bilingual learning attitudes will affect their inclination towards science knowledge and curriculum learning. The study also finds that Tibetan middle school students prefer Tibetan language learning, which has profound social reasons. For example, the development of Tibetan science learning models and mechanisms, the difficulty in understanding Chinese science knowledge, the preference of mother tongue and the prior knowledge Impact and so on.

Through the analysis of the three core elements of Tibetan science teachers, Tibetan science textbooks, and Tibetan science teaching process, the study found that there are irrational phenomena in the allocation of teacher resources in Tibetan middle schools, structural imbalances in the allocation of science teachers, and the phenomenon of cross-curricular teaching. The science teaching basic ability and extension ability of Tibetan science teachers need to be improved, the training mechanism of science teachers is in a passive situation, and information teaching literacy is lacking. At the same time, there are fewer experimental courses in the semester, and the experimental courses are mainly based on classroom demonstrations. The compiling and structural arrangement of science textbooks neglected the characteristics of cognitive development of Tibetan middle school students. Cases and materials lacked Tibetan cultural characteristics, and curriculum standards lack flexibility and science textbooks are more difficult. The study believes that the development of science education in Tibetan middle school is also plagued by socio-cultural ecological factors. On the ecological level of regional social and cultural atmosphere: the cultural atmosphere of the study of social sciences in Tibetan areas is weak, science knowledge is far away from the reality of life in Tibetan areas, the function of science education cannot be adapted to the actual needs of the Tibetan community, and at the same time, temple education has a strong impact on modern school education; in the family culture of education and education: the lack of family education for Tibetan middle school students, the traditional “heavy arts, light science” family education concept and gender education affect Tibetan students’ attitudes towards science education; in the aspect of individual cultural psychological ecology: Tibetan middle school students The unique national character will affect their tendency of science study; in the ecological aspect of the school system: the structural imbalance of the liberal arts and science major setting in the nationality colleges induces the biased tendency of the Tibetan high school liberal arts and sciences, and the restriction of the independent closed national education system. The way of science education for middle school students of Tibetans and the vicious circle of the career development of science teachers have reduced the teaching level of teachers. The active preferential

policies for college entrance examinations have not exerted practical effects on students' science studies. Therefore, the science education in the Gannan Tibetan area should implement the appropriate Tibetan-Chinese bilingual education model, that is, a hybrid Tibetan-Chinese bilingual education model, and strive to improve the professional quality and teaching ability of Tibetan science teachers and create an effective science classroom teaching model.

The science textbooks that characterize the development of students' cognition, at the same time, actively promote the cultural and ecological environment of Tibetan science education. On this basis, combined with the limitations of research, the future research trends are proposed.

Keyword: Tibetan, science, middle school students, cultural ecology

被默许的误认——当代大学教师对教育者身份的理解与建构的质性研究

曾妮 北京大学 博士

摘要:在高等教育质量提升日益紧迫的今天,大学教师的教育职责虽然屡被强调,却屡遭质疑。在高等教育大众化稳步推进的当下,大学教育的任务艰巨,但大学教师的精力常常分散至多处。究竟应该如何从大学教师之处着手,推动大学之教育功能的实现,这不仅是一个重要的研究课题,更是一个关键的实践命题。从已有研究来看,要么更多关注大学教师学术职业这一综合性的职业身份,要么更多采用思辨的方式指出大学教师职业伦理之应然状态,真正聚焦于大学教师教育者身份的研究寥寥可数。然而,如果不能有效解决“大学教师的教育者身份究竟具有何种内涵”这一前提性的问题,大学教师发展、大学教师职业道德等一系列问题的探讨,就很难有实质性的突破正是基于如上理由,本研究以“当代大学教师对教育者身份的理解与建构”为研究主题,通过对质性研究中扎根理论取径的应用,采用费尔克拉夫的批判话语分析理论对15名受访者的访谈文本进行了话语资源分析和话语秩序分析,以期能够通过鲜活的教育现场,揭示当前大学教师在教育者身份的自我理解与建构方面所呈出来的时代特征,从而有利于人们准确把握大学教师教育者身份的内涵,进而拓展大学教师职业认同研究的视阈,并为相关教育行政部门和大学管理者关于大学教师队伍建设的政策制定提供必要的事实依据。研究最重要的发现是:受访教师并非基于一种清晰的职业角色定位践行自己的教育者身份,而是根据自己的认识,“自由”选择教育者身份的定义。这种自由择直接表现为他们职业认同上的双重误读。而这种误读在中国大学的特殊情境中又很大程度上是被默许的。具体而言,研究的结论包括如下三方面:首先,受访教师们是通过内容、权责、意义和性质四个方面来描述自身的教育者身份的。其中,受访教师对于教育者身份性质的理解影响着其他几个方面的理解。根据教育者身份性质的两条属性(即教育者身份是否纳入到职业身份中,教育与研究是否相结合),受访者们可以分为六种类型:平行型、教支持研型、重研型、重教型、研支持教型和融通型。前三种类型的教师仅仅以制度所赋予的“大学教师”称谓来作为获得教育者身份的合法依据,其实质是一种以组织认同替代职业

认同的做法。而重教型和研支持教型的教师仅仅以个人对教育之一般属性的理解来完成对大学教师的教育者身份的全部描绘，这是一种以自我认同替换职业认同的行为。

其次，无论受访者是否存在职业认同上的误解，他们在建构自身教育者身份时，都使用了一些类似的话语资源，如市场话语、管理话语、知识话语、身份话语、同济话语和教育话语。只不过不同教师所构建的话语秩序有所不同：有的是市场置顶型，有的是管理置顶型，有的是教育置顶型。表面上看，话语资源的选择和话语秩序的组建似乎都是由教师个人意志所决定。但深入分析发现，大学教师对于自身教育者身份的误认，实际上是得到大学以及大学之外的中国社会情境的默许的。从大学的角度看，大学组织功能对职业内涵的替代性解读，大学知识特性对职业内涵的混淆式定位以及在此基础上存在的高等教育的教育迷失，都为大学教师对教育者身份认同的误认提供了中观的组织环境。从大学之外的市场和政治的角度看，中国社会的“官场+市场”的发展模式对大学教师的教育者身份建构提供了一种宏观的社会情境：强管理的模式使得教育工作化，教育管理的市场化导致教育的物质化，而学术的市场化也带来了教育的边缘化。这些因素的存在客观上也为大学教师的教育者身份误认提供了可能。

最后，既然大学教师的教育者身份存在误认，我们就需要重新确认这一身份。研究提出了确认的三重路径：1.职业身份的伦理建构，其根本目的是让大学教师回归本原性的教育者身份；2.组织认同的文化建构，旨在为大学教师践行本原性的教育者身份提供组织保障；3.自我认同的精神建构，重在唤醒大学教师的本真性教育者身份意识和意义感，从而为其教育者身份的有效落实提供强大的精神动力。

关键词：教育者身份，职业认同，教育与研究相结合，双重误认，本原性身份

Abstract: Nowadays, with the increasingly urgent to improve the quality of higher education, university teachers' education responsibilities have been repeatedly emphasized, but frequently doubted. With steady development of higher education massification, the mission of university education is arduous, whereas university teachers' energies are often expended on so many places. How should we enable the achievement of university's educational function by taking actions from university teachers? It is not only an important research topic, but also a key practical proposition. According to previous researches, researchers' research topics focus on either comprehensive professional identity of university teachers' academic profession, or the deserved state of their professional ethics in speculative ways. There are very few studies actually focus on university teachers' educator identity. However, if we cannot effectively solve the prerequisite issue which is "the connotation of university teachers' educator identity", it will be hard to achieve substantial breakthroughs on faculty development, university teachers' professional ethics, and so on. Based on the above reasons, this research selects "educator identity's understanding and construction of contemporary university teachers" as research subject. The researcher applies grounded theory and Fair clough's critical discourse analysis theory to analyze discourse resources and discourse structures within 15 interviewees' interview texts, hope to reveal the characteristics of contemporary university teachers' understanding and construction of educator identity through living education field study. The purpose of this study is to accurately grasp the connotation of university teachers' educator identity, expand research field of university teachers' occupational identity,

and provide necessary facts for relevant educational administrative departments and university administrators to make policies. The key finding of this research is that the interviewees do not practice their educator identity by clarified orientation of professional role, but by their own cognition to freely choose its definition. This freedom of choice leads to dual misidentification. In the particular context of Chinese university, misidentification is acquiesced. Specifically speaking, the conclusions of the study include the following three aspects:

First of all, the interviewees describe their educator identity through four aspects: content, right and duty, meaning, and nature. Among them, the teacher's understanding of the nature of educator identity influences their understanding of other aspects. According to the two attributes of educator identity's nature (Bringing educator identity to occupational identity; Combining education to research), the interviewees' educator identity could be divided into six types: Separated Understanding, 'Research Supported by Education' Understanding, Research-driven Understanding, Education-driven Understanding, 'Education Supported by Research' Understanding, and Harmony Understanding. The first three types of teachers only take "university teacher" appellation given by institution as the legal basis of educator identity, which is a way of replacing occupational identity with organizational identity. Teachers belong to type of Education-driven Understanding and 'Education supported by Research' Understanding only complete the description of university teachers' educator identity by personal understanding of education's general attribute, which is a way of replacing occupational identity with self-identity.

Secondly, whether or not the interviewees have misidentification of their occupational identity, they use similar discourse resources, such as market discourse, management discourse, knowledge discourse, status discourse, peer discourse and education discourse, to construct their own educator identity. Only the order of discourse constructed by different teachers differs. Some are market discourse set top, some are management discourse set top, and some are education discourse set top. On the surface, the choice of discourse resources and the construction of discourse structures seem to be determined by teachers' personal will. Analyzing deeply, however, it appears that university teachers' misidentification of educator identity is actually acquiesced by universities and other social situations in China. From universities' perspective, the alternative interpretation of the connotation of occupation by university organization function, the confused orientation of the connotation of occupation by knowledge attribute of university, and thus lost in higher education offer meso-level organization environment to the misidentification of university teachers' educator identity. From market and politics perspective, Chinese society's "Officialdom and Market" development pattern has built a macro-level social context for university teachers' construction of educator identity, which means strong management pattern has led education become a work; the marketization of education management has resulted in materialization of education, and the marketization of academics has rendered marginalization of education. These factors' existence also provides a possibility for misidentification of university teachers' educator identity.

Finally, since some teachers misidentify their educator identity, we need to reconfirm the identity. The study proposes three path of confirmation: 1. The fundamental goal of occupational identity' ethical construction is to make university teachers return to their original identity as educators; 2.the purpose of organizational identity's cultural construction is to provide organizational guarantee for university teachers to practice their original educator identity; 3.the focus of self-identification's spiritual construction is to awaken the sense of identity as authentic educators, thus providing a powerful spiritual impetus for the effective implementation of their educator identity.

Keyword: Educator Identity, Occupational Identity, Combination of education and research, Dual Misidentification, Original Identity

课堂教学公平及其指标体系研究

雷晓庆 南京师范大学 硕士

摘要:自人类社会诞生以来,公平公正便是人类社会发展的永恒价值追求,延伸到教育领域则表现为对教育公平的向往。随着义务教育的普及与发展,“教育起点公平”基本实现以后,“教育过程公平”逐步成为推进教育公平的关键。课堂教学作为最基本的教育过程活动,是教育过程公平的重要组成部分,既反映教育过程公平,又左右教育过程公平,其公平状况理应受到重视。然而,课堂教学公平涉及到教育公平中较为微观的问题,很难从较为权威的数据库中找到相关描述性数据。因此,在借鉴已有相关研究的基础上,本研究试图以指数研究的方式对课堂教学公平问题进行探究。

本文主要遵循“分析——建构——解释——验证”的基本思路,运用文献分析法、德尔菲专家咨询法、问卷调查法等方法,分阶段对课堂教学公平内涵及其指标体系进行研究,全文共分为五部分内容。

第一部分主要是通过文献的查找与阅读,梳理和分析了国内外与课堂教学公平及其指标体系相关性较强的研究,并指出课堂教学公平指标体系的建构要立足于课堂教学公平实际问题、关注课堂教学的整体性、注重主客观因素的双重作用。

第二部分主要是对课堂教学公平的内涵进行了深入解析。从人的双重属性——共同的人性与独特的个性出发,结合个人对课堂教学公平的认知,指出课堂教学公平就是在课堂教学过程中,教师要以平等对待为底线,平等尊重每个学生,根据学生的特征及其发展水平,不同对待不同学生,并对弱势群体进行适当的补偿对待,但不以损害精英群体的发展为代价,让平凡者不平凡,优秀者更优秀,使每个学生都能得其应得,发展成为最好的自己。

在第三部分中,本研究首先明确了课堂教学公平指标体系的建构依据,在遵从一定的建构原则的前提下,设置了5个一级指标、13个二级指标、39个三级指标

以及包括教学特权、教学歧视、教学排斥、教学羞辱在内的具体指标，初步建构了课堂教学公平指标体系；随后，通过三轮德尔菲专家问卷的发放与修改，最终形成了由 4 个一级指标，14 个二级指标，27 个三级指标构成的课堂教学公平指标体系。

第四部分主要是对最终建构的课堂教学公平指标体系框架进行了简要介绍，厘清了各级指标间的关系，并且根据专家们对各级各类指标的排序，运用矩阵对偶法，确定了各级各类指标权重，明晰了各项指标在指标体系中所占的比重；随后，对课堂教学公平指标体系的内容进行解释说明，并制定了课堂教学公平指标标准，既增强了对课堂教学公平指标体系的认知，也增强了课堂教学公平指标体系的可操作性。

在第五部分中，为检测和提高已建构的课堂教学公平指标体系的效度，本研究设计了课堂教学公平测量问卷，并选取 N 市 R 中学初二年级的 5 个班级 207 名学生进行了问卷的试测，之后对其进行了项目分析，确定了问卷题项；然后，对 J 省 S 市和 N 市两所学校同一年级 10 个班级的学生进行了问卷的正式发放，利用 SPSS22.0 统计软件对调查问卷进行了分析，发现：校级之间课堂教学公平存有明显的差异、班级之间课堂教学公平也存有明显的差异等等。因此，通过小范围的应用，初步断定本研究所建构的课堂教学公平指标体系既能够测得出公平，也能推得进公平。

关键词：课堂教学公平 指标 指标体系

Abstract: Since the birth of human society, fairness and justice have been the eternal value pursuit of social development, which extending to the education filed has been the yearning for education fairness. With the popularization and development of compulsory education, "education starting point fairness" has been basically realized, and "education process fairness" has gradually become the key to promote education fairness. As the most basic activity in education process, classroom teaching is a component part of education process fairness, which can not only reflect the education process fairness, but also control the education process fairness. Thus, the fair condition of classroom teaching should be taken seriously. However, classroom teaching fairness involves some relatively micro problems of education fairness, so it is difficult to find relevant descriptive data from the authoritative database. Therefore, based on the relevant researches, this study attempts to explore classroom teaching fairness in the way of index research.

This study follows the basic idea of "analysis--construction—interpretation

--verification", and applies to some methods such as literature analysis, Delphi expert consultation and questionnaire investigation, in order to study the connotation and index system of classroom teaching fairness phase by phase. This paper is divided into five parts.

The first part mainly analyzes the researches strongly related with index system of classroom teaching fairness both at home and abroad through literature searching and reading, and points out that the construction of index system of classroom teaching fairness should base on the practical problems of classroom teaching fairness, and pay attention to the integrity of classroom teaching and the dual functions of subjective and objective factors.

The second part is mainly about the connotation of classroom teaching fairness. Started with human's dual attributes--human generality and individual uniqueness and combined with individual's perception of classroom teaching fairness, it is pointed out that classroom teaching fairness means that in the process of classroom teaching, teachers should take equal treatment as the bottom line, and respect each student equally. Moreover, teachers should treat students differently according to their characteristics and development levels, and give some appropriate compensation to the disadvantaged group without the detriment to the elite group's development, in order to let the ordinary be extraordinary, the good be better, and make every student "get his deserved" to become the best.

In the third part, first of all, this study has been clear about the construction basis of index system of classroom teaching fairness. Secondly, under the premise of following the certain construction principles, this study has preliminary constructed an index system of classroom teaching fairness including 5 first-level indicators, 13 second-level indicators, 39 third-level indicators and some specific indicators such as teaching privileges, teaching discrimination, teaching rejection and teaching humiliation. Then, through the three rounds of distribution and modification of Delphi expert questionnaire, an index system of classroom teaching fairness including 4 first-level indicators, 14 second-level indicators and 27 third-level indicators has been finally formed.

The fourth part mainly introduces the framework of constructed index system of classroom teaching fairness, and clarifies the relationship between indicators at all levels. Then, according to the experts' sorting of various indicators at all levels, the weight of each indicator in the index system has been determined by using the matrix

dual method. Next, explaining the content of the index system of classroom teaching fairness, and formulating the standard of classroom teaching equity index, which not only strengthens the cognition of classroom teaching equity index system, but also enhances the operability of classroom teaching equity index system.

In the fifth part, this study designed a survey questionnaire of classroom teaching fairness to test and improve the validity of constructed index system of classroom teaching fairness. A total of 207 students from 5 classes of grade 2 of R middle school in N city were selected to conduct the questionnaire test. After the project analysis, the questionnaire items were finally determined. Then, the questionnaire was formally sent to the students from 10 classes in the same grade of S city and N city in J province, and analyzed by the statistical software SPSS22.0. It is found that classroom teaching fairness has significant differences between schools and classes. Therefore, through a small range of applications, it is preliminary concluded that the index system of classroom teaching fairness that constructed by this study not only can measure the fairness, but also can promote the fairness.

Keyword: Classroom teaching; Teaching fairness; Indicators; Index system