

美国教育评价研究的知识基础与热点

——基于 2000—2014 年 8 种 SSCI 期刊的可视化分析

邱均平, 欧玉芳

摘要: 美国教育评价由心理测量占统治地位的“测量时代”发展而来。通过可视化分析, 可知 21 世纪美国教育评价的知识基础是建立在教育学、统计学、行为科学和心理学的学科知识之上, 尤其借鉴了统计学和心理学的研究方法。研究热点在理论上集中于项目反应理论、概化理论和经典测试理论; 在研究方法上对基于模型的定量化评价非常重视, 其中罗殊模型和 5-因素模型是美国教育评价研究领域的常用研究模型; 在研究内容上主要关注问责制、标准制定、学生成绩、引导程序、动机、抑郁症以及情商等评价问题的研究。

关键词: 美国教育评价; 知识基础; 研究热点; 可视化分析

中图分类号: G649.712 文献标识码: A 文章编号: 1671-0169(2016)02-0142-08

DOI:10.16493/j.cnki.42-1627/c.2016.02.016

20 世纪以来, 美国教育评价研究经历了三个时代: 以“测量”理论的形成和测验技术在实际中的大量运用为标志的“测量时代”, 凭“测验结果”作出描述为特征的“描述时代”, 以“价值判断”作为评价理论为主要特色的“判断时代”^{[1](P1-6)}。进入 21 世纪后, 其研究发生了什么变化? 基于哪些知识基础? 热点研究领域是什么? 本文使用文献计量软件 Bibexcel 和可视化软件 CiteSpace II 3.8.R1, 通过文献计量和可视化技术, 阐释 21 世纪以来美国教育评价研究的研究热点和知识基础, 以期为相关研究提供有价值的参考。

一、数据来源和研究方法

(一) 数据来源

笔者从 2014 年 SSCI 期刊目录的标题中以“evaluation”(评价)、“assessment”(评估)和“measurement”(测量)查找所有期刊, 查找后删除非教育评价的期刊, 最后选取了 8 种专门研究美国教育评价的同行评审期刊作为数据来源。据 JCR 报告, 这 8 种期刊的 5 年影响因子都在 1 以上, 可以说他们具有较好的代表性, 所选期刊如表 1 所示。

(二) 研究方法

选定研究对象后, 笔者开始获取研究对象。首先在 Web of Science (WOS) 数据库中以 8 种美国教育评价期刊的名称为“American Journal of Evaluation” or “Applied Psychological Measurement” or “Educational Evaluation and Policy Analysis” or “Educational and Psychological Measure-

表 1 美国 8 种 SSCI 教育评价期刊目录及简介

期刊名称	5 年影响因子	发行频次	选取文章数
《美国评价杂志》 (American Journal of Evaluation)	1.527	季刊	388
《教育评估、评价和绩效》 (Educational Assessment Evaluation and Accountability)	1.153	双月刊	144
《教育评价和政策分析》 (Educational Evaluation and Policy Analysis)	2.685	季刊	282
《教育和心理测量》 (Educational and Psychological Measurement)	1.582	双月刊	851
《评价评述》 (Evaluation Review)	1.276	双月刊	351
《国际评选和评估》 (International Journal of Selection and Assessment)	1.042	季刊	511
《教育测量杂志》 (Journal of Educational Measurement)	1.064	季刊	298
《应用心理测量》 (Applied Psychological Measurement)	1.153	双月刊	465
选取文章数小计			3288

ment” or “Evaluation Review” or “International Journal of Selection and Assessment” or “Journal of Educational Measurement” 检索式, 检索类型选择“出版物”, 检索时间设定为 2000 年至 2014 年, 在检索结果中选择“article”(研究型文章)和“proceedings”(会议论文)精简检索结果, 得到 3288 条包含文章标题、关键词、摘要以及参考文献的记录。然后, 用词频统计软件 Bibexcel 统计 8 种期刊的高频关键词。最后, 用可视化软件 CiteSpace II 3.8.R1, 并借助其文献共被引聚类、关键词共现和突变词检测功能, 从多变角度反映美国教育评价研究的研究热点及知识基础^[2]。具体研究过程如图 1 所示。

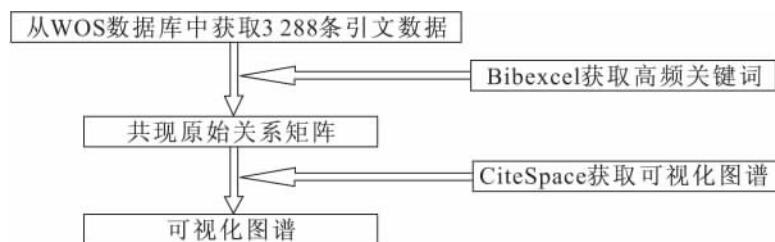


图 1 研究对象可视化处理流程

二、知识基础分析

在文献计量学中, 施引文献形成研究前沿, 被引文献构成知识基础^[3]。在众多的被引文献中高被引文献最能体现研究领域的知识基础^[4], 因此, 本研究通过分析美国教育评价研究的高被引文献和关键节点文献来透视其知识基础。分析美国教育评价的知识基础可使研究人员更好地把握美国教育评价的发展脉络和研究基础。知识基础能反映研究前沿本质, 美国德雷塞尔大学陈超美教授开发的 CiteSpace II 3.8.R1 中的共被引聚类功能可以很好地体现研究领域的知识基础。设置 CiteSpace II 3.8.R1 的相关参数, 并选择图谱显示类型为聚类视图 (Cluster View), 运行软件得到 21 世纪以来美

国教育评价研究领域的高被引文献可视化知识图谱(如图2所示)。

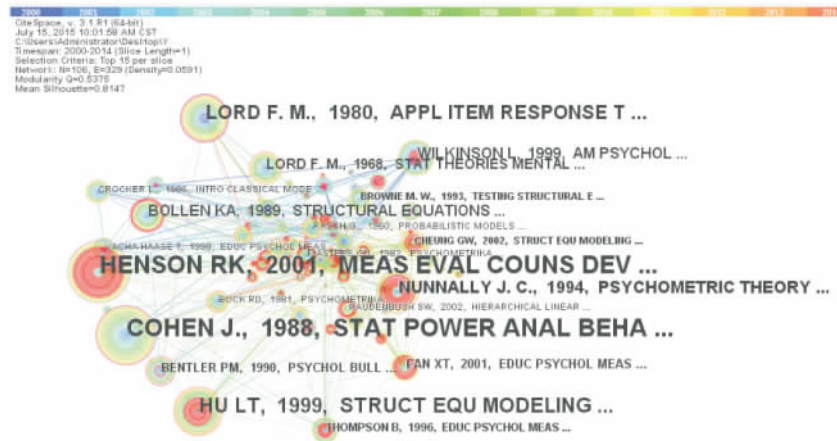


图2 21世纪以来美国教育评价研究领域高被引文献可视化图谱

图2节点的大小与节点相对应的文献被引频次成正比,节点越大表明该文献的被引次数越高。选取被引频次 ≥ 55 的文献作为21世纪以来美国教育评价研究领域的高被引文献,通过对文献被引频次高低进行分析后发现21世纪以来美国教育评价研究领域共有7篇高被引文献。这7篇高被引文献依次如下:第一篇是2001年罗宾·汉森(R. K. Henson)的《理解信度估计的内部一致性:阿尔法系数的概念内涵》^[5],该研究对教育评价最常用的信度估计、内部一致性系数和阿尔法系数进行了研究,并为教育评价研究者们提出了一套理解以上参数的解释性概念框架。第二篇是1988年美国统计学家雅各布·科恩(Jacob Cohen)的统计学著作《行为科学的统计强度分析》^{[6](P1-20)},该著作以一种应用的视角提供了一种独特的统计概率分析,它包含12章内容,分别是强度分析概念、均值T检验、积差 r_s 的显著性、相关系数的显著性、比例为0.50的检验以及符号检验、比例的显著性、拟合度卡方检验和相依表、方差分析和协方差分析、多元回归和相关分析、集合相关和多因素分析方法、强度分析中的常见问题以及计算过程。第三篇是1980年由有着“现代测试之父”之称的弗雷德里克·洛德(Frederick M. Lord)的著作《项目反应理论在实际测试问题中的应用》^{[7](P1-8)},此书分为四部分,第一部分简要评述了经典测试理论,并介绍了几个潜在特质模型及其概念、假设;第二部分和第三部分主要阐述项目反应理论在实际测试问题中的应用,包括项目参数在试题和试题库研发中的使用、判断测试对不同能力水平的被试者的有效性以及两种测试的有效性的比较、构建新测试和重新设计旧测试、标准参照测试的设计和评价、等值分数以及试题偏差的研究;第四部分洛德阐述了估计真分数分布的一些新理念。第四篇是1999年李子胡(Li-tze Hu)和彼得·本特勒(Peter M. Bentler)的《协方差结构分析中的拟合指数中断标准:传统标准与新标准的比较》^[8],此文研究了用于判定拟合指数的传统中断标准和几个新标准的适切性,其中拟合指数是用来评价模型在实践中的拟合程度的。第五篇是1994年由农纳利·久姆(Nunnally Jum)和伯恩斯坦·艾勒(Bernstein Ira)编著的《心理测量理论》^{[9](P1-16)},该著作是美国教育学和心理学研究生教材,它研究了教育学和心理学中的测量问题,将经典统计推理程序与现代统计推理程序相结合。第六篇是利兰德·威尔金森(Leland Wilkinson)和统计推断小组撰写的《心理学期刊使用的统计方法:指南和解释》^[10],该文针对1994年雅各布·科恩在《美国心理学家》上发表的《地球是圆的》(*The Earth Is Round*)一文引起有关心理学期刊使用显著性检验的争议,对心理学研究的方法设计、抽样、测量参数、工具、程序、强度和样本大小做了解释。第七篇是肯尼思·博伦(Kenneth A. Bollen)的著作《结构方程潜在变量》^{[11](P1-36)},此书证明了结构方程模型的通用性,强调路径分析、递推模型、非递推模型、古典经济计量学以及验证性因素分析的应用,并探索了建模过程中实质性知识发挥的重要作用。

由此可见, 21 世纪以来美国教育评价研究领域的知识基础建立在统计学、行为科学和心理学的学科知识之上, 尤其借鉴了统计学和心理学学科的研究方法, 例如结构方程模型、路径分析、递推模型、验证性因素分析等量化研究方法被大量地借鉴到教育评价学中。这正好反映了教育评价学是一门综合学科的学科属性。

三、研究热点分析

(一) 美国教育评价研究高频关键词分析

关键词指从文献中选取的反映文献主题, 满足检索需要的词或词组。这不多的词或词组概括文献或归纳文献的核心内容, 或揭示文献的主题信息, 或提炼文献的观点和论证方法^[12]。作为学术论文不可或缺的组成部分, 尽管关键词所占篇幅极少 (通常不过 3 至 8 个词), 它们却是论文的精髓^[13]。分析文献关键词可得出前沿的科学研究热点。文献计量学通常根据关键词的词频及其共现网络图来对其进行分析。词频分析指提取文献的关键词, 统计其频次, 确定高频关键词, 由高频关键词分析研究对象的热点、前沿和发展趋势的一种分析方法^[14]。笔者用 Bibexcel 软件统计 3288 篇美国教育评价研究文献后, 得到了这些文献的关键词及其词频, 限于篇幅, 本研究仅选取词频大于或等于 10 的关键词 (如表 3 所示)。

表 3 美国教育评价研究高频关键词 (频次 ≥ 10 的关键词)

序号	高频关键词	频次	序号	高频关键词	频次
1	item response theory (项目反应理论)	126	21	student achievement (学生成绩)	17
2	confirmatory factor analysis (验证性因素分析)	68	22	bootstrap (引导程序)	16
3	reliability (信度)	68	23	exploratory factor analysis (探索性因素分析)	15
4	validity (效度)	63	24	measurement error (测量误差)	14
5	factor analysis (因子分析)	51	25	logistic regression (逻辑回归)	14
6	differential item functioning (项目功能差异)	49	26	missing data (缺失数据)	14
7	Rasch model (罗殊模型)	43	27	generalizability theory (概化理论)	13
8	meta-analysis (元分析)	31	28	depression (抑郁症)	13
9	effect size (效应量)	30	29	assessment (评估)	13
10	measurement invariance (恒等性)	27	30	equating (均衡)	12
11	measurement (测量)	26	31	multidimensionality (多维度)	12
12	reliability generalization (信度概化)	25	32	coefficient alpha (阿尔法系数)	12
13	structural equation modeling (结构方程建模)	23	33	unidimensionality (单向度)	11
14	confidence interval 置信区间	23	34	dif (离差值)	11
15	computerized adaptive testing (计算机化自适应测试)	20	35	parallel analysis (平行分析)	11
16	nonparametric item response theory (非参数项目反应理论)	19	36	measurement equivalence (测量对等性)	10
17	factor structure (因子结构)	18	37	emotional intelligence (情商)	10
18	accountability (问责制)	18	38	type I error (第一类错误)	10
19	construct validity (结构效度)	18	39	kernel smoothing (核平滑估计)	10
20	standard setting (标准设定)	18	40	classical test theory (经典测试理论)	10

由表 3 可知, 在研究理论层面, 21 世纪美国教育评价主要集中研究三大理论, 即: 项目反应理论 (频次 145, 包括频次为 126 的项目反应理论和频次为 19 的非参数项目反应理论)、概化理论 (频次 13) 以及经典测试理论 (频次 10), 这三大理论都产生于教育评价发展的初期阶段, 即忽略价值判断的教育测量阶段, 因此, 有学者视其为测量理论^{[15](P21)}。经典测量理论是历史上的第一个, 同时也是最一般、最基本的测量理论, 经过长期的发展, 现在较为成熟。但是经典理论有其自身无法克服的缺点, 如项目参数对样本的依赖性, 测量误差分析的笼统性等。概化理论和项目反应理论是为克服经

典测量理论的缺陷而发展起来的现代测量理论,这两个新理论尽管仍在发展与完善之中,但已经显示出强大的生命力,尤其是项目反应理论,其出现频次高达 145 次,位居关键词频次的首位,已成为 21 世纪美国教育评价的主流研究理论。

在研究方法层面,21 世纪美国教育评价研究主要使用了验证性因素分析(频次 68)、信度分析(频次 68)、效度分析(频次 63)、因子分析(频次 51)、元分析(频次 31)、置信区间分析(频次 23)、平行分析(频次 10)和核平滑估计分析(频次 10)等。由于西方的教育评价方法的发展过程主要是由标准化测验产生到扬弃的过程^{[16](P47)},因此 21 世纪美国教育评价研究方法倚重信/效度分析、因子分析以及元分析等教育测量研究方法是理所当然的。我们可以认为 21 世纪美国教育评价研究方法是在继承传统研究方法的基础上向现代先进的教育评价方法(验证性因素分析、平行分析和核平滑估计分析)发展。

在研究内容层面,21 世纪美国教育评价研究主要注重问责制(频次 18)、标准设定(频次 18)、学生成绩(频次 17)、引导程序(频次 17)、动机(频次 15)、抑郁症(频次 13)以及情商(频次 10)等评价问题的研究。

(二) 美国教育评价研究热点可视化图谱分析

词共现网络分析也叫共词分析,是一种文本内容分析技术,它通过分析在同一个文本主题中的款目对(单词或名词短语对)共同出现的形式,确认文本所代表的学科领域中相关主题的关系,进而探索学科的发展^[17]。我们可以通过关键词共现可视化图谱来确定研究主题或者研究热点之间的关系。使用 CiteSpace II 3.8.R1 的聚类分析功能,以对数似然函数比例算法(Log-likelihood Ratio, LLR 算法),设置适当阈值,从关键词中提取名词来表征聚类标签后,笔者得到了基于关键词的可视化聚类图谱,图谱共 99 个节点,95 条连线(如图 3 所示)。

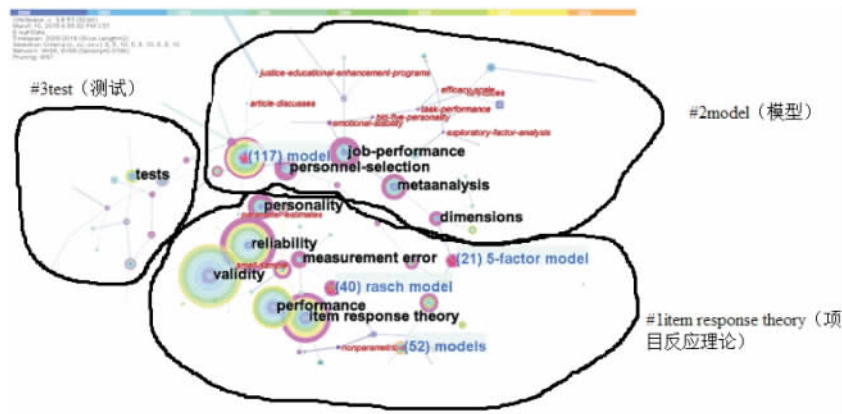


图 3 21 世纪以来美国教育评价研究热点可视化图谱

图 3 显然形成了 3 大聚类,分别是 #1item response theory (项目反应理论)、#2model (模型)和 #3test (测试)。其中, #1item response theory (项目反应理论)聚类中由 performance (绩效)、rasch model (罗殊模型)、validity (效度)、reliability (信度)、measurement error (测量误差)、small-sample (小样本)以及 nonparametric-item response theory (非参数项目反应理论)等关键词构成; #2model (模型)由 justice-educational-enhancement-programs (教育公平促进项目)、article-discusses (文章论述)、emotional-stability (情绪稳定性)、task-performance (工作绩效)、efficacy-scale (效力量表)、exploratory-factor-analysis (试探性因子分析)、job-performance (工作绩效)、personnel-selection (人员甄选)、metaanalysis (元分析)以及 dimensions (维度)等高频关键词构成;而 #3test (测试)聚类是个小聚类,由一些与测试相关的关键词构成,由于这些关键词与其他关键词共现次数少,图谱未显示具体名称,故而不对其进行研究。

此外,图谱中 model/ models (模型,其中 model 与其他关键词的共现次数为 117 次,models 与其他关键词的共现次数为 52 次)、rasch model (罗殊模型,与其他关键词的共现次数为 40 次)和 5-factor model (5-因素模型,与其他关键词的共现次数为 21 次)与相关的关键词关系密切。因此,我们可以看到美国教育评价对基于模型的定量化评价非常重视。

罗殊模型又称罗殊分析法,最初由丹麦数学家乔治·罗殊 (Georg Rasch) 于 1953 年提出,它是一种分析分类数据的心理计量模型,可用来权衡被试者的能力、态度或者人体特点与项目难度的关系^{[18](P3)}。其中,除罗殊提出的二分罗殊模型之外,还包括等级量表模型、部分计分模型、测验等值化等多个重要分析方法。因此,除了应用于心理计量学和教育研究领域之外,具有普适性的罗殊模型及其扩展模型也可用于其他领域,包括卫生和市场研究领域^[19]。在 21 世纪的美国教育评价研究中,许多研究^[20-28]以罗殊模型为主要工具对学生考试数据进行等值化处理,然后基于处理数据结果对考试的试题和学生的学业表现做出评价,以实现考试评价和对评价解释的标准化。

1981 年戈德堡 (Goldberg) 在人格特征论先驱 Norman 的研究方法和结果基础上提出了 5-因素模型,它是一种用来描述人格特征的心理模型,其中五个心理因素指神经质 (Neuroticism)、外倾性 (Extraversion)、经验开放性 (Openness to Experience)、愉悦 (Agreeableness) 和尽责性 (Conscientiousness),也被称为“大五” (Big Five) 模型^[29]。早在 20 世纪末,美国心理学研究人员就从事了 5-因素模型与教育方面的研究,例如: Ehrler 等人经研究证明:无学习问题的儿童人格特征多呈现为愉悦、情绪稳定、好辩论和不懈努力;相反,学习问题较多的儿童常具有退缩以及冲动的人格特征。他们发现愉悦、谨慎特征、与开放性特征较低的学生容易具有社会问题、处理不良问题、注意力障碍以及过动症征兆。此外,他们也发现情绪稳定和审慎特征人格特征均与学科成绩表现有显著相关^[30]。21 世纪以来,从事美国教育评价研究的学者们主要集中于 5-因素模型与教育或学校(如学业成就、阅读偏好等方面)的相关性及其在教育评价中的作用方面的研究。例如: Ridgell 等人证明人格特征与学业成就有显著相关关系^[31]。Schutte 和 Malouff 以 251 位大学生为研究样本,研究阅读偏好与人格的关系,结果显示五大人格特质可以显著预测出阅读偏好,开放性特质与尽责特质是阅读偏好良好的预测指标,整体而言,5-因素模型可以显著预测阅读的数量及质量^[32]。由此可见,21 世纪以来美国教育评价研究大量借鉴了的心理学研究方法和工具(如:罗殊模型和 5-因素模型),这与教育评价发展历史的三阶段,即固定考试型时期、心理测量占统治地位时期以及后现代时期^{[33](P1-10)},是密不可分的。纵然 21 世纪的美国教育评价属于后现代时期的教育评价,但它是由心理测量占统治地位时期的“测量时代”发展而来的,因此,21 世纪的美国教育评价研究大量借鉴心理学研究方法和工具也就不足为奇。

四、结果分析

本文以美国科学情报研究所编辑出版的 2014 年社会科学版期刊引证报告为依据,选取 8 种高影响力外文期刊(五年影响因子 ≥ 1.0) 21 世纪以来(2000—2014 年)所刊载的 3 288 篇研究论文及其题录数据作为研究样本,利用信息可视化的计量研究方法,从多方面对美国教育评价研究领域的知识基础和研究热点进行分析与展示,得出如下结论:第一,通过 2000—2014 年美国教育评价研究领域高被引文献共现网络知识图谱的绘制,发现 21 世纪以来美国教育评价研究领域的知识基础建立在教育学、统计学、行为科学和心理学的学科知识之上,尤其借鉴了统计学和心理学学科的研究方法。美国教育评价研究领域的知识基础研究结果充分证明了教育评价学是一门综合学科的学科属性,实践中的教育评价需要教育学、管理学、心理学、数学、统计学和经济学等越来越多的综合性学科知识来支持。第二,通过绘制 2000—2014 年美国教育评价研究领域关键词聚类视图,得出 21 世纪以来美国教育评价研究领域的研究热点在研究内容上集中探索三大理论,即:项目反应理论、概化理论和经典测试理论;在研究方法上对基于模型的定量化评价非常重视,其中罗殊模型和 5-因素模型是美国教育

评价研究领域的常用研究模型。此外,具体研究方法又从信/效度分析、因子分析以及元分析等教育测量研究方法向验证性因素分析、平行分析和核平滑估计分析等现代先进的教育评价方法发展的趋势;最后,在研究内容上主要关注问责制、标准设定、学生成绩、引导程序、动机、抑郁症以及情商等评价问题的研究。

参考文献

- [1] [美]埃贡·G·古贝,[美]伊冯娜·S·林肯.第四代评估[M].秦霖,等,译.北京:中国人民大学出版社,2008.
- [2] Chen,C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature [J]. *Journal of the American Society for Information Science and Technology*,2006,(3).
- [3] Persson,O. The intellectual base and research fronts of JASIS 1986—1990[J]. *Journal of the American Society for Information Science*,1994,(1).
- [4] 邱均平,吕红.近五年国际图书情报学研究热点、前沿及其知识基础——基于17种外文期刊知识图谱的可视化分析[J]. *图书情报知识*,2013,(3).
- [5] Henson,R. K. Understanding internal consistency reliability estimates; A conceptual primer on coefficient alpha [J]. *Measurement and Evaluation in Counseling and Development*,2000,(3).
- [6] Cohen,J. *Statistical Power Analysis for the Behavioral Sciences*[M]. Florence:Routledge,1988.
- [7] Lord,F. M. *Applications of Item Response Theory to Practical Testing Problems*[M]. New Jersey:Lawrence Erlbaum Associates,1980.
- [8] Hu,L. ,P. M. Bentler. Cutoff criteria for fit indexes in covariance structure analysis; Conventional criteria versus new alternatives[J]. *Structural Equation Modeling—A Multidisciplinary Journal*,1999,(1).
- [9] Nunnally,J. ,I. Bernstein. *Psychometric Theory* (3rd Ed.)[M]. New York:McGraw Hill,1994.
- [10]Wilkinson,L. ,Task Force on Statistical Inference. Statistical methods in psychology journals; Guidelines and explanations[J]. *American Psychologist*,1999,54.
- [11]Bollen,K. A. *Structural Equations with Latent Variables*[M]. New York:Wiley—Interscience,1989.
- [12]孙旭,程学竹.网络学习社区的自组织形态研究[J]. *中国电化教育*,2006,(12).
- [13]黄伟.教师网络学习社区的被组织和自组织[J]. *中国远程教育*,2011,(163).
- [14]欧玉芳.21世纪以来美国教师教育研究[D].金华:浙江师范大学,2013.
- [15]顾海根.人员测评[M].合肥:中国科技大学出版社,2005.
- [16]张远增.高等教育评价方法研究[M].上海:复旦大学出版社,2002.
- [17]姜春林,等.经济学研究热点领域知识图谱:共词分析视角[J]. *情报杂志*,2008,(9).
- [18]Rasch,G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Expanded Edition with Foreword and Afterword by B. D. Wright[M]. Chicago:The University of Chicago Press,1980.
- [19]Wright,B. D. Solving measurement problems with the rasch model[J]. *Journal of Educational Measurement*,1977,(2).
- [20]Verguts,T. ,P. D. Boeck. A rasch model for detecting learning while solving an intelligence test[J]. *Applied Psychological Measurement*,2000,(2).
- [21]Lin,P. Y. ,Y. C. Lin. Examining student factors in sources of setting accommodation DIF[J]. *Educational and Psychological Measurement*,2014,(5).
- [22]Williams,R. T. ,A. Swanlund,et al. Measuring instructional differentiation in a large-scale experiment[J]. *Educational and Psychological Measurement*,2014,(2).
- [23]Van,K. -S. ,R. R. Meijer. Detection of person misfit in computerized adaptive tests with polytomous items[J]. *Applied Psychological Measurement*,2002,(2).
- [24]Kaliski,P. K. ,S. A. Wind,et al. Using the many-faceted rasch model to evaluate standard setting judgments: An illustration with the advanced placement environmental science exam[J]. *Educational and Psychological*

Measurement,2013,(3).

- [25]Adams,R. J. ,M. L. Wu,et al. The rasch rating model and the disordered threshold controversy[J]. *Educational and Psychological Measurement*,2012,(4).
- [26]Jiao,H. ,J. H. Liu,et al. Comparison between dichotomous and polytomous scoring of innovative items in a large-scale computerized adaptive test[J]. *Educational and Psychological Measurement*,2012,(3).
- [27]Koning,E. D. ,K. Sijtsma,et al. Comparison of four IRT models when analyzing two tests for inductive reasoning[J]. *Applied Psychological Measurement*,2002,(3).
- [28]Digman,J. M. Personality structure: Emergence of the five-factor model[J]. *Annual Review of Psychology*,1990,(41).
- [29]Goldberg,L. R. An alternative “description of personality”: The big-five factor structure[J]. *Journal of Personality & Social Psychology*,1990,(6).
- [30]Ehrler,D. J. ,J. G. Evans. Extending big-five theory into childhood: A preliminary investigation into the relationship between big-five personality traits and behavior problems in children[J]. *Psychology in the Schools*,1999,(6).
- [31]Ridgell,S. D. ,J. W. Lounsbury. Predicting academic success: General intelligence, “big five” personality traits, and work drive[J]. *College Student Journal*,2004,(4).
- [32]Schutte,N. S. ,J. M. Malouff. University student reading preferences in relation to the big five personality dimensions[J]. *Reading Psychology*,2004,(4).
- [33]Lewy,A. *International Encyclopedia of Curriculum*[M]. Oxford:Pergamon Press Ltd. ,1991.

(责任编辑 孙洁)