

我国语文教学表现性评价的热点研究

——基于知识图谱可视化分析

关怀怀, 齐雪艳

伊犁师范大学中国语言文学学院, 新疆 伊宁

收稿日期: 2024年3月25日; 录用日期: 2024年4月23日; 发布日期: 2024年4月29日

摘要

为了准确了解我国语文教学中表现性评价研究热点, 采用关键词共词分析法, 运用Bicomb2.0及SPSS26.0软件, 以中国知网为资源平台, 对国内的440篇文献进行可视化分析。研究发现: 我国语文教学表现性评价已经发展起来了, 但是在教师表现性评价能力方面有所欠缺、核心素养和深度学习融合有待发展、表现性评价的评价标准和准则不够规范等一系列问题。

关键词

语文教学, 表现性评价, 可视化分析

A Hotspot Study on Performance Evaluation of Chinese Language Teaching in China

—Visualization Analysis Based on Knowledge Graph

Huaihui Guan, Xueyan Qi

Chinese Language and Literature College of Yili Normal University, Yining Xinjiang

Received: Mar. 25th, 2024; accepted: Apr. 23rd, 2024; published: Apr. 29th, 2024

Abstract

In order to accurately understand the research hotspots of performance evaluation in Chinese language teaching in China, the keyword co-occurrence analysis method was adopted, and Bicomb2.0 and SPSS26.0 software were used. China National Knowledge Infrastructure (CNKI) was used as the resource platform to visually analyze 440 domestic literature. The research results have found

文章引用: 关怀怀, 齐雪艳. 我国语文教学表现性评价的热点研究[J]. 教育进展, 2024, 14(4): 1048-1054.

DOI: 10.12677/ae.2024.144625

that performance evaluation in Chinese language teaching has developed in China, but there is a lack of teacher performance evaluation ability, the integration of core competencies and deep learning needs to be developed, and the evaluation standards and guidelines for performance evaluation are not standardized enough.

Keywords

Chinese Language Teaching, Performance Assessment, Visualization Analysis

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

美国对表现性评价提出了描述性定义：表现性评价是要求学生创造出答案或产品以展示其知识或技能的测验[1]。这是在 20 世纪 90 年代后成为欧美各国及我国香港地区的教育界和测量界的热门课题，它的理论是以建构主义为背景，以原有知识为基础，根据具体情境，建构用于指导问题解决的模式，而不能靠将已有知识简单提取出去解决实际问题[2]。之所以受到欢迎的原因是人们对于测验性评价的日渐不满。人们认为这样的测量题只能对简单知识进行衡量，只能了解学生知道什么，而不知道能做什么。语文作为我国的基础学科，工具性与人文性都极为重要，表现性评价既可以测量知识掌握情况，也可以与学生思想素养做互动。它的出现让教 - 学 - 评一体化，也推动了语文素养的提升。

知识图谱以科学为基础，涉及应用数学、信息科学及计算机科学诸学科交叉的领域，是科学计量学和信息计量学融合的新发展，以可视化的图像直观地展现出来[3]。所以利用知识图谱可以把研究我国表现性评价这一领域的相关文献作为抓手，借助科学的数据和方法，凭借真实的资料与结论，发现该领域研究的热点问题及发展趋势。

2. 资料来源与研究方法

2.1. 资料来源

以中国知网数据库为资料来源，从旧版入口进入，进行高级检索，将表现性评价和语文作为关键词进行检索。共检索到期刊文献 856 篇。对文献进行取舍，将不符合要求的外文文献，报纸，会议、无作者信息、刊物总目录信息、戏曲评论、影评等剔除点，最终选择 440 篇作为研究资料。最后规范资料，检查资料编码格式以及关键词含义是否一致。

2.2. 研究工具

以 Bicomb 共词分析软件和 SPSS26.0 为主要研究工具。其中 Bicomb 共词分析软件由中国医科大学医学信息学系崔雷教授和沈阳市弘盛计算机技术有限公司开发。Bicomb 基本操作步骤为：首先，选取研究资料。研究对象是 CNKI 所检索出的 440 篇文献，共涉及关键词 749 个。其次，收集和整理资料。将收集到的 CNKI 格式数据库文献转化为 Bicomb 能够识别的 ANSI 编码文本文件。第三，运用 Bicomb 软件进行关键词统计。在该软件中建立新项目，然后选取该项目格式为 cnki 默认格式，接着选取已经整理好的资料所在目录；提取该文档进行关键词统计，最后进行关键词共词矩阵分析导出共词矩阵。

2.3. 研究进程

首先, 确定关键词。采用 Bicomb 软件对 440 篇文章进行关键词统计, 抽取出词频大于 5 的 27 个词作为主要关键词。其次, 建立高频关键词共词矩阵。对 27 个关键词利用 Bicomb 软件中的共词分析功能生出共词矩阵。再次, 进行聚类分析。将共词矩阵导入 SPSS26.0 中, 采用样本聚类, 得出关键词聚类树图。最后, 结合聚类和研究热点图谱内容进行解释和分析。

3. 研究结果与分析

3.1. 高频关键词及分析

利用 Bicomb2.0 软件对 440 篇文献进行关键词统计, 选取出排名前 27 位的关键词为高频关键词, 结果见表 1。可以看出语文表现性评价研究较热门的前 27 个热点排名词分别为表现性评价、核心素养、实验教学、美国、表现性任务、形成性评价、应用、评价、深度学习、教师表现性评价、教师评价、教学评价、教学、教学设计、学业评价、职前教师、评价标准、评价量表、表现性评定、发展性评价、师范生、项目式学习、音乐教学、档案袋评价、教师专业发展、评价量规、教师。仅对高频关键词统计难以发现它们间的联系, 还需要进一步通过关键词共词技术来深入挖掘它们之间的联系。见表 1:

Table 1. Table of keyword frequency for performance evaluation of Chinese language teaching

表 1. 语文教学表现性评价关键词频次表

序号	关键词	频次	序号	关键词	频次	序号	关键词	频次
1	表现性评价	348	11	教师评价	7	21	师范生	5
2	核心素养	35	12	教学评价	7	22	项目式学习	5
3	实验教学	16	13	教学	7	23	音乐教学	5
4	美国	14	14	教学设计	6	24	档案袋评价	5
5	表现性任务	13	15	学业评价	6	25	教师专业发展	5
6	形成性评价	12	16	职前教师	6	26	评价量规	5
7	应用	10	17	评价标准	6	27	教师	5
8	评价	9	18	评价量表	6			
9	深度学习	8	19	表现性评定	6			
10	教师表现性评价	7	20	发展性评价	6	合计		580

3.2. 高频关键词的 Ochiai 系数相似矩阵分析

为了得到符合尺度分析要求的数据结构, 将上述 27 个高频关键词汇合利用 Bicomb 软件进行共词分析生成词篇矩阵。将词篇矩阵导入 SPSS26.0, 选取 Ochiai 系数将其转化为一个共词相似矩阵, 该矩阵仅选取一部分结果见表 2。

表 2 中的数字表明数据间的相异性。数字的大小表明相应两个关键词之间的距离远近, 其数值越接近 1, 说明关键词之间的距离越近、相似度越大; 反之, 数值越接近 0, 关键词之间的距离越大、相似度越小。从表 2 可以看出表现性评价与核心素养之间的关系是很近的, 这说明表现性评价的发展对核心素养的提升密切相关, 同时也可以看到表现性评价与美国、实验教学、表现性任务、形成性评价、应用、深度学习、教师评价等距离也都比较近。而与教师表现性评价比较远, 说明表现性评价主要运用于学生身上, 是对学生设计任务的。这些距离远近说明, 在已发表的研究文献中, 核心素养、实验教学、表现性任务结合起来论述的研究成果多于将其他几种关键词结合的研究成果。

Table 2. Similarity matrix for Chinese language performance evaluation (partial only)**表 2.** 语文表现性评价相似矩阵(仅部分)

个案	表现性评价	核心素养	实验教学	美国	表现性任务	形成性评价	应用	评价	深度学习	教师表现性评价	教师评价
表现性评价	1.000	0.290	0.188	0.143	0.119	0.093	0.170	0.018	0.133	0.000	0.081
核心素养	0.290	1.000	0.085	0.000	0.000	0.000	0.000	0.169	0.000	0.000	0.000
实验教学	0.188	0.085	1.000	0.000	0.000	0.000	0.000	0.083	0.000	0.000	0.000
美国	0.143	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.101	0.000
表现性任务	0.199	0.000	0.000	0.000	1.000	0.000	0.000	0.092	0.000	0.105	0.000
形成性评价	0.093	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.102	0.000	0.000
应用	0.170	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
评价	0.018	0.169	0.083	0.000	0.092	0.000	0.000	1.000	0.000	0.000	0.000
深度学习	0.133	0.000	0.000	0.000	0.000	0.102	0.000	0.000	1.000	0.000	0.000
教师表现性评价	0.000	0.000	0.000	0.101	0.105	0.000	0.000	0.000	0.000	1.000	0.286
教师评价	0.081	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.286	1.000

3.3. 生成高频关键词聚类图

利用 SPSS26.0 对表 2 的高频关键词相似矩阵进行聚类分析, 聚类结果可直观地表明高频关键词间的远近关系, 结果见图 1 所示。图中纵轴数字代表与之相应的高频关键词, 横轴数字代表关键词之间的距离, 如果在更短的距离(横轴方向)内, 两个关键词聚集在一起, 说明二者相关度更高, 关系更密切; 反之则相关度低[4]。

从图 1 可以看出我国语文表现性评价研究高频关键词可以分为六类。种类一: 对于教师的表现性评价: 主要包括教师评价、教师专业发展、教师表现性评价、职前教师四个关键词。直观反映出表现性评价需要教师的直接操作, 在真正的运用之前要先将教师自身的素养提高, 自身理解与把握表现性评价的操作与定义。种类二: 表现性评价在教学上的应用。它被引进到教学应用中, 说明在教育实践过程中是值得引用和借鉴的。种类三: 是对表现性评价的界定与要求, 包括关键词发展性评价、档案袋评价、形成性评价深度学习。表现性评价在教育中也是可以看为形成性评价和发展性评价的, 它们的本质与应用是相通的, 也可以被称为替代性评价和真实性评价[5]。同时它的出现也促进了深度学习的进程。种类四: 关于表现性评价的设定, 它包括的关键词有教学设计、评价量规、表现性任务、项目式学习、核心素养、表现性评价、实验教学、评价、教学评价、评价标准。它们之间的距离很近, 表现性评价作为一种评价方式, 指向于核心素养, 置于评价的任务, 并敲定了评价标准。它与实验教学不同, 最初可以理解为在项目式学习中的一种评价, 带有评价量规和准则[6]。种类五: 对它的起源和适用人群介绍, 最早提出与运用是在美国, 教师作为主导学生依旧是教学主体, 在他们之间用表现性评价加强互动与联系。种类六: 对表现性评定的研究, 可以把它用来对学业进行评定。这种评定直观反映学生学业表现, 是不同于纸笔测验的评价。

3.4. 关键词多维尺度分析

为进一步探寻关键词之间隐藏的关键, 利用 SPSS26.0 对 27 个关键词构成的相似矩阵进行多维尺度分析, 标准化方法采用 Z 分数, 得到我国语文教学表现性评价研究热点的可视化图谱, 见图 2 所示。

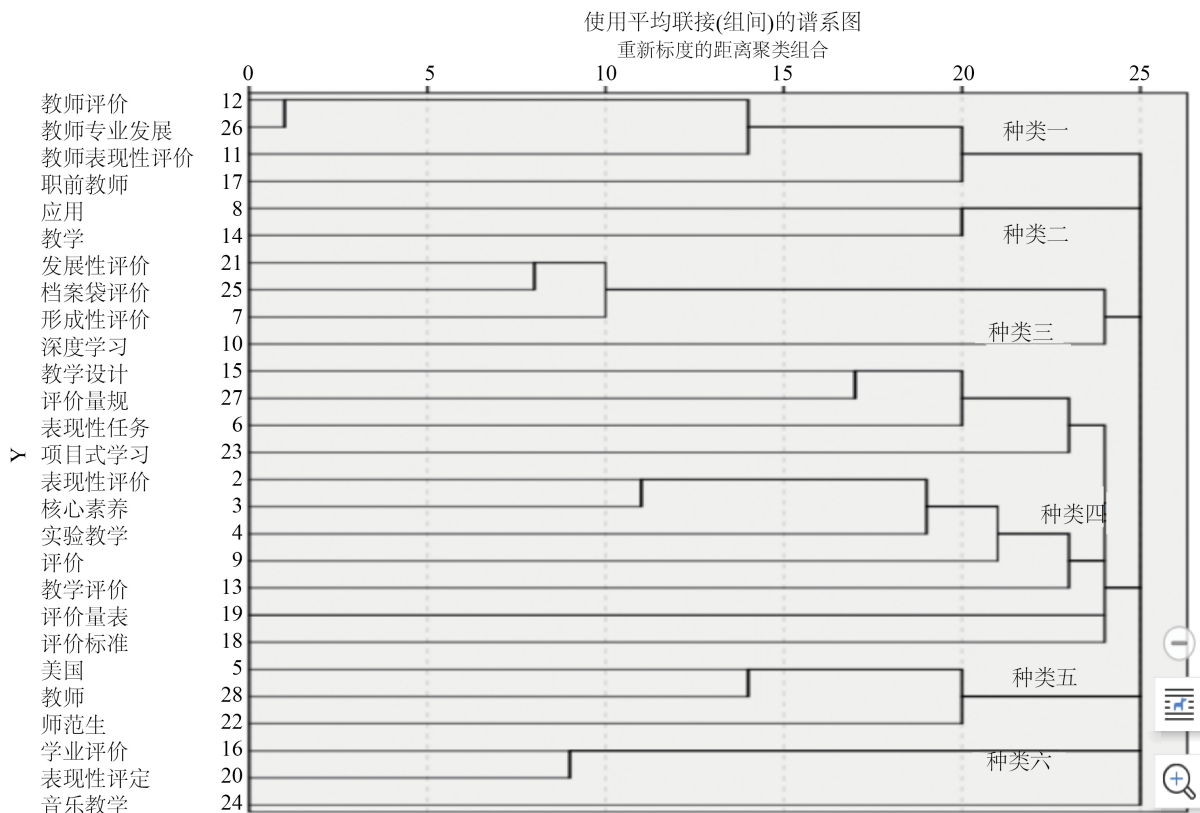


Figure 1. Keyword clustering diagram for Chinese language performance evaluation
图 1. 语文表现性评价关键词聚类图

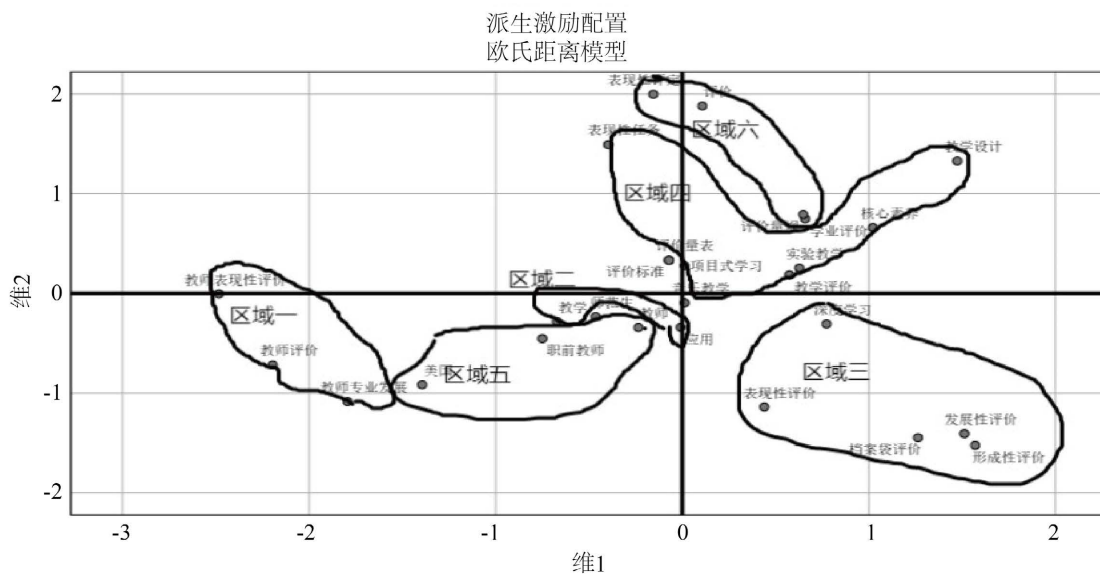


Figure 2. Multidimensional analysis knowledge graph of Chinese language performance evaluation
图 2. 语文表现性评价多维分析知识图谱

将关键词进行多维尺度分析绘成知识图谱，通过分析关键词在坐标位置间的距离，来判断它们之间联系疏密的程度[7]。图 2 中坐标称为战略坐标，关键词的位置由小圆圈来表示，关键联系紧密的程度则

由小圆圈之间的距离来判断, 关键词联系紧密则代表他们的小圆圈距离近, 反之, 距离远[8]。圆圈距离坐标中心越近, 其所代表的关键词影响力越大。

以逆时针方向将图2划分为四个象限, 将关键词划分为六个区域。区域一中的教师表现性评价位于横坐标, 其余两个都在第三象限, 说明现在的研究是教师进行的表现性评价, 且在未来对教师的评价会趋向于对教师的专业发展, 有很大的研究潜力。区域二中的师范生、应用分别位于横坐标和纵坐标, 在我国表现性评价已广泛的用于师范生的教育教学, 它们之间的联系非常紧密的。区域三中发展性评价和形成性评价几乎重叠, 它们的定义和运用都说明, 表现性评价是可以和它们互为替代, 本身意义也相通, 只是研究不深。其中深度学习靠近横坐标也说明, 深度学习的发展是值得融合的方向。区域四中的评价量表和评价标准以及教学都在坐标轴上且靠近中心坐标, 并且区域四大部分在第一象限, 这更能说明我国的语文表现性评价很热门, 研究比例增大, 侧重对核心素养的发展。不断被用于教学设计和学业评价之中。区域五的位于第三象限, 此时美国推出在职前对教师进行表现性训练, 未来我们也可以借鉴把研究重点放在在职前教师的专业发展, 让表现性评价发挥更大优势。区域六中表现性评定和评价距离很近, 它们已经被频繁研究且有很大的成果, 未来可以将重心放到别的领域。

4. 我国语文教学表现性评价研究趋势

4.1. 教师表现性评价的能力需要重点关注和提升

构成区域一的关键词在第三象限且联系不紧密, 距离坐标也很远, 但是其中的教师表现性评价是距离坐标最近的。说明在这一区域中, 教师的影响因素是最大的, 它揭示了教师在语文教学表现性评价中的作用愈来愈重要。教师是课程实践的主体, 教师是否掌握表现性评价的基本原理、程序及要求, 是否能在实践中主动、有效地应用表现性评价, 是影响评价改革, 乃至整个课程改革向纵深发展的关键[9]。为了加强我国教师表现性评价的训练, 具体建议有: 第一在职教师的培训要采用“在做中学”的思路, 边培训, 边实践, 使教师切实掌握设计、分析、反思和改进; 第二改进师范生教育。师范生在校期间要加强包括表现性评价在内的各种实践技能的培养, 改革师范生课程设置, 将其作为必修课程并且更新课程内容。

4.2. 核心素养的养成和深度学习的发展需要融合

区域四中的核心素养处于第一象限, 在区域中和教学设计、学业评价、教学评价联系很紧。区域三中的深度学习靠近坐标轴, 它和核心素养一样都是可以和表现性评价进行有效融合的。在深度学习中加强表现性评价有利于核心素养的养成与发展, 这就是它们三者之间的关系。核心素养是个体适应未来社会生存与发展所需要的关键能力、必备品格与价值观念[10]。它超越了简单的知识记忆和技能训练, 指向真实情境中的问题解决, 同样深度学习要求学生在学的过程中, 教师可以不断引导让她们学会不断反思与批判。说到这里我们就可以发现, 它们的要求和表现性评价的展开都是有相似之处的, 所以如何把三者之间的关系平衡好就需要下一番功夫了。

4.3. 表现性评价的评定标准应该得到规范和落实

区域四中的评价量表位于坐标轴上, 距离坐标中心也很近。另外, 评价量规和学业评价几乎重合。说明了表现性评价作为一种评价方式, 是有别于量表的质性评价, 它的评价准则与制定都是很重要的。对于表现性评价的评价准则和要求, 杨丽提出可以将教学目标设置成可观测性的, 表现性任务创设为有情境的, 对于准则建构易操作性的[11]。在表现性评价中, 将目标设置可观测的可以让学生的学习活动对目标有个清晰的了解, 并且在真实的情境中进行任务, 让教师可以就学生的具体表现给出评价, 这样也

能学生明白将评价置于真实并做出反应。最后, 将标准可操作化并且多维角度的构建, 学生可以对评分细则有个明确的了解, 也清楚自己该怎么行动能取的好的成绩。目前我国的现状是没有明确如何规范设置评价标准, 让教师和学生都没有可具体参考的参照物。

5. 结语

采用科学图谱法可直观地反映出我国语文教学表现性评价的研究主要围绕六大领域展开。可视化的统计研究结果表明, 我国语文教学表现性评价的研究相较于外国来说还是比较缺乏, 且涉猎领域不多。不过表现性评价在教学中的应用还是较多的, 也在慢慢往核心素养的方向发展。但是有一系列问题的暴露: 教师表现性评价能力缺失、核心素养和深度学习的融合没有得到足够的发展、评定标准不够规范等等。所以我国语文教学表现性评价的研究还有待进一步的深化与拓展。

本研究利用研究热点知识图谱的方法将量化的数据可视化, 既形象又不失严谨。但根据关键词共词矩阵进行分析, 区分不了跨领域研究和重复研究, 再加上聚类分析结果和知识图谱存在部分差异, 因此需要研究者借鉴更多的先进方法进行后续研究。

参考文献

- [1] (美)肯尼斯·莫尔. 课堂教学技巧[M]. 刘静, 译. 北京: 人民教育出版社, 2010.
- [2] 黄玲玲. 浅谈表现性评价在语文教学中的应用[J]. 福建广播电视大学学报, 2013(3): 68-72.
- [3] 陈悦, 刘泽渊. 悄然兴起的科学知识图谱[J]. 科学学研究, 2005, 23(2): 149-154.
- [4] 郭文斌, 张晨琛. 我国融合教育热点领域及发展趋势研究[J]. 残疾人研究, 2017(3): 63-69.
- [5] 赵德成. 表现性评价: 历史、实践及未来[J]. 课程·教材·教法, 2013, 33(2): 97-103.
- [6] W. James Popham, 著. 促进教学的课堂评价[M]. 国家基础教育课程改革“促进教师发展与学生成长的评价研究”项目组, 译. 北京: 中国轻工业出版社, 2003: 142.
- [7] 郭文斌. 知识图谱理论在教育与心理研究中的应用[M]. 杭州: 浙江大学出版社, 2015: 88-107.
- [8] 郭文斌, 周念丽, 方俊明. 我国学前教育研究热点知识图谱[J]. 学前教育研究, 2014(1): 11-18.
- [9] 周文叶. 中小学表现性评价的理论与技术[M]. 上海: 华东师范大学出版社, 2021: 53-54.
- [10] 崔允漷, 邵朝友. 试论核心素养的课程意义[J]. 全球教育展望, 2017, 46(10): 24-33.
- [11] 杨丽. 深度学习下表现性评价的内涵、特点及实施策略[J]. 语文建设, 2023(2): 59-64.