

# 复杂性理论视角的学术评价理论建构 ——基于扎根理论的探索性研究

陆红如<sup>1</sup>, 阮选敏<sup>1</sup>, 成颖<sup>1,2</sup>, 陈雅<sup>1</sup>

(1. 南京大学信息管理学院, 南京 210046; 2. 山东师范大学文学院, 济南 250014)

**摘 要** 学术评价问题一直受到广泛关注, 近年来颁布的各项政策法规表明了国家对该问题的高度重视。本文对中英文学术评价理论和应用研究及代表性的评价实践进行了扎根分析。通过开放编码归纳出 18 个初始评价要素; 主轴编码形成 8 个关键评价要素; 在选择编码过程中通过复杂性理论基本原理和关键要素的指引, 确定了关键评价要素之间的逻辑关系, 完成了学术评价理论的建构。扎根分析的结果显示, 在评价语境的影响下, 评价主体在与评价客体及其利益相关者相互作用中形成明确的评价思想(评价目的、评价分类、评价维度、理论基础), 并进一步确定评价条件(评价标准、评价制度、质量控制), 对评价实施进行约束和规范。形成评价思想、确定评价条件和进行评价实施的过程中产生实时评价产出(评价数据)和最终评价产出(评价结果)。

**关键词** 复杂性理论; 扎根理论; 学术评价; 评价实践

## Construction of an Academic Evaluation Theory from the Perspective of Complexity Theory—Exploratory Research Based on Grounded Theory

Lu Hongru<sup>1</sup>, Ruan Xuanmin<sup>1</sup>, Cheng Ying<sup>1,2</sup> and Chen Ya<sup>1</sup>

(1. School of Information Management of Nanjing University, Nanjing 210046;

2. School of Chinese Language and Literature of Shandong Normal University, Jinan 250014)

**Abstract:** Evaluation in academia is an issue of wide concern in society. The policies and regulations promulgated in recent years have shown that China attaches great importance to this issue. This paper employs grounded theory to analyze research on evaluation theory and practice and summarizes 18 primary evaluational factors through open coding and eight key evaluational factors through axial coding. Finally, selective coding determines the logical relationship between key evaluational factors, guided by the basic principles and pivotal elements of complexity theory, to construct a theory of evaluation in academia. The analysis shows that based on the evaluation context, the evaluator forms clear evaluation thoughts (evaluation purpose, evaluation classification, evaluation dimension, and theoretical basis) through interaction with the evaluation object and stakeholders, and further determines evaluation conditions (evaluation criteria, evaluation system, and quality control) to restrict and standardize the implementation of evaluation. Real-time evaluation output (evaluation data) and final evaluation output (evaluation result) are generated through the overall evaluation process.

**Key words:** complexity theory; grounded theory; evaluation in academia; evaluation practice

收稿日期: 2019-07-25; 修回日期: 2019-12-08

基金项目: 国家社会科学基金项目“施引者引用意向与文献计量视角的学术论文被引影响因素研究”(17BTQ014)。

作者简介: 陆红如, 女, 1993 年生, 博士研究生, 主要研究方向为信息行为; 阮选敏, 女, 1994 年生, 硕士研究生, 主要研究方向为信息计量; 成颖, 男, 1971 年生, 博士, 教授, 博士生导师, 主要研究方向为用户信息行为、信息检索, E-mail: chengy@nju.edu.cn; 陈雅, 女, 1965 年生, 博士, 教授, 博士生导师, 主要研究方向为图书馆管理与服务。

## 1 引言

学术评价问题一直以来受到社会的广泛关注,近期密集发布的多项政策(如2018年7月中共中央办公厅和国务院办公厅印发的《关于深化项目评审、人才评价、机构评估改革的意见》,2018年10月科技部、教育部、人力资源社会保障部、中科院和中国工程院联合发布的《关于开展清理“唯论文、唯职称、唯学历、唯奖项”专项行动的通知》等)一方面可以看出国家对学术评价问题的高度重视,另一方面也从侧面反映了我国学术评价领域存在一些亟待解决的问题。

马克思主义认为,“理论的价值主要在于理论对实践具有巨大的指导作用。要使理论对实践产生指导作用,使实践达到预期效果,就必然要求确保理论的科学性”<sup>[1]</sup>。习近平总书记在十九大报告中更强调了理论发展的重要性,指出“实践没有止境,理论创新也没有止境”,“我们必须在理论上跟上时代,不断认识规律,不断推进理论创新、实践创新、制度创新、文化创新以及其他各方面创新”<sup>[2]</sup>。由此可见,要解决学术评价领域存在的问题必须重视科学的学术评价理论的建构和应用。学术评价理论作为该领域的知识基础,带动着整个领域的发展,为评价主体和客体提供沟通交流的系统化语言,为评价活动的开展和评价问题的解释提供宏观指引及微观描述。

目前,国内外学者就学术评价理论已经形成了较丰硕的研究成果,包括学术评价理论的建构<sup>[3]</sup>、评析<sup>[4]</sup>和应用<sup>[5]</sup>。从已有研究成果来看,学者们构建的评价理论大多针对特定学术领域,尚未见应用广泛的学术评价理论。此外,学者们针对同一内涵的概念使用的术语不尽相同。例如,文庭孝<sup>[6]</sup>的工作中使用了“评价规范体系”等,而吕红<sup>[7]</sup>的工作则使用了“评价标准体系”等。鉴于此,本文采用扎根理论进行普适性学术评价理论建构的探索性研究。学术评价活动具有高度复杂性,评价主体需要基于不同的评价目的在具体的评价语境之下准确识别评价问题,制定出全面、真实、客观地反映评价客体的评价指标体系。与此同时,评价客体的状态随着时间的推移在不断发生变化,相应的评价活动也需要做出及时的动态性调整。构建用于指导和解释学术评价活动的理论有必要充分考虑评价活动的复杂性,故“复杂性理论”即成为自然之选,将为本研究的扎根分析提供方法论框架,该理论中的整

体性、系统性和自组织性思想将贯穿于学术评价理论建构的整个过程。

## 2 文献述评

### 2.1 评价理论的含义及分类

美国著名哲学家约翰·杜威认为,评价本质上是一种价值判断,起到引导行为的作用<sup>[8]</sup>。评价作为一个科学领域,需要有不断发展的评价理论作为支撑。对于评价理论的含义,学者们有不同观点。例如,Shadish<sup>[9]</sup>认为评价理论揭示了评价的本质,提供了讨论评价问题的交流语言;Donaldson等<sup>[10]</sup>提到了评价理论应该包含的几个部分:知识(产生可靠知识的方法)、使用(如何利用知识)、价值(如何构成价值判断)和实践(评价者在现实世界环境中如何实施评价);Hansen等<sup>[11]</sup>总结出评价理论囊括的几个要素,即假设(提示或促进评价方法的潜在信念)、评价情境(评价发生的环境,可包括场景、需求、资源等)、评价活动(用于规划和开展评价的程序)、评价结果/效果(评价及实施过程中可能发生的变化,可能对个人和团体等产生的影响或变化)和外部因素(不受评价影响但可能影响评价的条件或事件)。

学界根据不同的维度对评价理论进行了分类。①根据评价理论的适用范围,Shadish<sup>[9]</sup>将其分为面向特定领域和多个领域2种类型,另外还有一类基于经验的假设性理论正逐渐发展成为评价理论的新分支。②根据评价理论的性质,Scriven<sup>[12]</sup>将其分为规范性和描述性,规范性理论确定了评价的内涵,描述性理论则阐释了评价类型、评价实践及机理等。③根据评价理论的内容,Alkin等<sup>[13]</sup>将其分为方法、价值判断和使用三类,第一类主要关注评价研究中涉及的方法论,第二类强调评价主体进行价值判断的过程,第三类专注于评价数据的使用者和使用方式。

### 2.2 学术评价理论与实践研究

#### 2.2.1 学术评价理论的梳理

目前,学术评价理论集中于教育、人才、成果、项目、机构、人文社科、科研管理等领域。具体而言,教育领域内,学者们主要从教育资源利用效率、教育装备、教育收益等角度进行评价理论的建构或梳理<sup>[14-17]</sup>。例如,辛涛等<sup>[4]</sup>将教育评价划分为科学取向、管理取向、人类学取向和参与取向4

种，并阐述了相应的教育评价模式；Bergsmann等<sup>[18]</sup>分理想情况、实际情况和教学过程3个阶段阐述了高等教育机构能力本位教学的评价理论。众多学者基于不同的理论进行了人才评价研究。例如，王媛等<sup>[19]</sup>借鉴模糊理论研究了人才评价方法；赵伟等<sup>[20]</sup>基于胜任力模型理论与个体创新行为理论提出了包括创新知识在内等6个方面的创新型科技人才评价冰山模型。对于成果评价，许海云等<sup>[21]</sup>考虑了技术类知识产品价值测度中的不确定性问题，构建出核心价值评价指标体系并确定了相应的权重。关于项目和机构评价，乔欢<sup>[22]</sup>从构建、环境、尺度和方法4个角度分析了Saracevic的数字图书馆评价理论框架；Millstone等<sup>[23]</sup>借鉴变化理论，从利益相关者参与度等角度探讨了科技创新项目的监测和评估方法的优化方式。在人文社科和科研管理领域，影响力较大的有叶继元<sup>[3]</sup>面向人文社会科学领域构建出的包括形式评价、内容评价和效用评价3个维度，以及评价主体、客体、目的、标准及指标、方法和制度六大要素的“全评价”理论。文庭孝<sup>[6]</sup>面向科研管理工作中的科学评价，从主体体系、内容体系、规范体系、理论体系、活动体系方面构建了科学评价理论体系。

### 2.2.2 学术评价实践案例

本文选择世界范围内影响较大的学术评价实践作为案例，从实践角度进行评价理论的梳理。世界四大权威机构（《美国新闻和世界报道》U.S. News<sup>[24]</sup>，泰晤士高等教育Times Higher Education<sup>[25]</sup>，英国高等教育资讯和分析数据提供商Quacquarelli Symonds<sup>[26]</sup>，上海软科教育信息咨询有限公司Shanghai Ranking Consultancy<sup>[27]</sup>）的最新版全球大学排名评价报告的方法部分公布了评价方法和评价指标的选择和依据，未提及学术评价理论。由4个英国高等教育资助机构（英格兰研究、苏格兰资助委员会、威尔士高等教育资助委员会和北爱尔兰经济部）共同组成的基金会，致力于评价英国高等教育机构的研究质量，在其提出的评价框架REF（research excellence framework）的REF2014中将研究的广泛影响力、成果质量和研究环境的活跃程度纳入评价范围，但是也未明确其依据的学术评价理论<sup>[28]</sup>。由南京大学中国社会科学评价中心公布的《中文社会科学引文索引》来源期刊/集刊遴选实施方案（试行）中提出了“定量（文献计量指标）评价与定性（学科专家）评价相结合”等评价原则，也未见相应的学术评价理论的阐述<sup>[29]</sup>。上海社会科学院

智库研究中心《2017中国智库报告》中提到了评价标准、原则、类别、程序和结果，未见提及明确的评价理论依据<sup>[30]</sup>。上述学术评价实践也有可能借鉴了评价理论，但未在报告中明确指出；无论是哪种情况都可以看出目前学术评价理论未体现其应有的位置。

## 2.3 复杂性理论在评价领域的应用

近年来复杂性理论（complexity theory）在评价领域得到了越来越广泛的讨论与应用。Walton<sup>[31]</sup>总结了46篇学术期刊文献中关于复杂性理论在评价中的运用，提炼出8个维度，包括形成对系统的理解，吸引子、涌现性和其他复杂性思考，定义特定层级和分析单元，评价的时机，参与式评价，案例分析和比较设计，多重和综合性方法，以及指导评价的层级理论。Jorm等<sup>[32]</sup>将复杂性理论运用到医学院评价中，认为当前的医学院评价没有考虑学习与教学、研究、社区参与等方面的关系，忽略了更广泛的系统产生的影响，在研究中提出需要综合考虑复杂性理论的嵌套、多样性、自组织性和涌现性结果等特征。Eoyang等<sup>[33]</sup>认为越来越复杂的评价语境对传统的评价方法和工具提出了挑战，需借鉴复杂性理论理解评价实践中的动态性和协同作用。Walton<sup>[34]</sup>认为在评价中使用复杂性理论面临着机遇和挑战，虽然复杂性理论可以帮助解决评价中存在的问题，但是复杂性的多重定义、现有评价方法的主导地位等制约了复杂性理论的应用。目前，复杂性理论在评价领域内的运用多以实践为主，较少有运用复杂性理论进行评价理论的建构研究。

## 2.4 述评

国内外学术评价理论的相关研究经过几十年的发展，已经覆盖越来越多的领域，囊括了评价的各方面内容，但是目前缺少普适性学术评价理论的建构研究。多数学者明确限定了其构建的评价理论面向的特定领域，且部分学者针对同一领域根据不同的标准、原则和价值观构建了多个相异的学术评价理论体系，导致学术评价实践很难在大量的研究成果中找到最适合的评价理论，也限制了学术评价理论的实际应用。普适性学术评价理论的缺乏导致评价理论的应用范围过于狭窄、经验可转移性较低，难以指导学术评价的科学有序开展。通过分析全球范围内影响力较大的学术评价实践可知，目前评价实践中对于评价理论的应用还不够，所以进行普适



性学术评价理论的建构研究来推动其在评价实践中的应用,扩大学术评价理论的影响力极为必要。鉴于复杂性理论在评价实践中研究与应用的广泛性以及取得的显著成效,其可以为普适性学术评价理论的建构提供研究框架。

### 3 研究设计

#### 3.1 研究方法

本研究采用扎根理论展开研究,通过开放编码和主轴编码从评价理论和评价实践相关研究成果中解析出初始评价要素(范畴)和关键评价要素(主范畴)。在选择编码阶段,借鉴复杂性理论中的基本原理,通过对关键评价要素的抽象提炼出核心范畴,探索关键评价要素间的逻辑关系和内在机理,构建出学术评价理论模型。研究设计流程如图1所示。

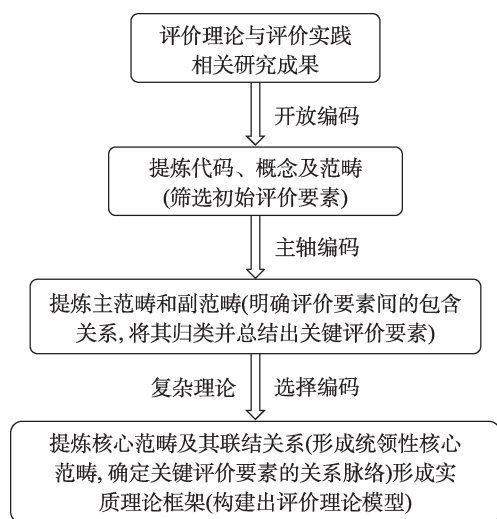


图1 研究设计流程图

#### 3.2 理论抽样和资料收集

评价的理论与实践之间存在着密切的联系。评价作为一个以实践为导向的领域,关于评价实践的经验知识对于发展相关的评价理论起到关键作用<sup>[35]</sup>。实际开展评价时面临的各类困难催生了评价模型,新的问题、环境和参与者也促使评价理论得到不断的完善。在构建评价理论时应关注评价实践中的微观层面,分析出对评价实践产生影响的各类因素,并且通过分析评价活动中的各类偶然性可以总结出外在多样性中潜在的内在统一性。发展评价理论不仅要从各成熟理论中吸收精华,更需要从实践中总结经验,以便更好地指导实践。评价理论的构建离不开对评价理论和评价实践相关研究成果的

深入调研。据此,本文的资料来源是中英文学术评价理论和实践的相关研究成果。为保证研究资料的质量,本文选择发表在SCI、SSCI、CSSCI或CSCD来源期刊上的学术评价理论和实践研究论文以及国内外影响较大的评价应用。

##### 3.2.1 研究资料的搜集

###### 3.2.1.1 国内学术评价理论研究的搜集

在CNKI数据库中以来源期刊=CSSCI或CSCD期刊,主题=评价理论或评估理论为检索条件,得到1429条结果;以主题=评价体系或评估体系,且主题=理论,得到1349条结果。对以上检索结果以2条标准进行筛选:①创新型研究成果,即排除综述类、比较类、书评类和应用类;②剔除与本研究主题无关的语言学中的“评价理论”和心理学中的“认知评价理论”。通过上述2条排除标准及去重处理,共得到289条结果。阅读其题名和摘要,共筛选出81篇学术评价理论相关论文。

###### 3.2.1.2 国外学术评价理论研究的搜集

在Web of Science (WoS)核心合集数据库中进行检索,检索条件为: Title= ("assess\*" OR "evaluat\*" OR "apprais\*" OR "measur\*" OR "valu\*" OR "estimat\*" OR "rat\*") AND ("theory" OR "framework") AND ("teacher" OR "student" OR "project" OR "publication" OR "academ" OR "scien\*" OR "research" OR "education" OR "journal" OR "universit\*" OR "intelligen\*"),选择SCI和SSCI引文索引库,得到1279条检索结果。阅读题名和摘要之后,参照国内学术评价理论的筛选标准,共获得82篇学术评价理论相关论文。

###### 3.2.1.3 国内学术评价实践的搜集

在CNKI数据库中,以“主题=评价实践,来源期刊=CSSCI或CSCD期刊”为检索条件进行检索,结果为765条。通过浏览其题名和摘要,筛选出72篇学术评价实践相关论文。除了论文以外,国内学术界影响范围较大的评价实践有南京大学中国社会科学评价中心的《中文社会科学引文索引》来源期刊/集刊遴选活动、上海社会科学院智库研究中心的全国性智库评价、上海软科教育信息咨询有限公司的全球大学排名等。

###### 3.2.1.4 国外学术评价实践的搜集

在WoS核心合集数据库中进行检索,检索策略为: Title= ("assess\*" OR "evaluat\*" OR "apprais\*" OR "measur\*" OR "valu\*" OR "estimat\*" OR "rat\*") AND ("practice") AND ("teacher" OR "student" OR "project"

OR "publication" OR "academ" OR "scien\*" OR "research" OR "education" OR "journal" OR "universit\*" OR "intelligen\*"), 选择SCI和SSCI引文索引库, 得到2227条结果。浏览检索结果后, 筛选出106篇学术评价实践相关论文。

除了论文以外, 国外影响力较大的学术评价实践还有美国宾夕法尼亚大学“智库研究项目”(TTCSP)的全球智库评价, 《美国新闻与世界报道》、《泰晤士高等教育》、英国高等教育资讯和分析数据提供商Quacquarelli Symonds的全球大学排名评价等。

### 3.2.2 研究资料的选择

根据本研究学术评价理论建构的目的, 采取以下标准筛选样本: ①领域的全面性: 由题名和摘要的阅读可知学术评价理论和实践覆盖教育、机构、成果、人员、项目、人文社科、科研管理等领域, 进行资料的选择时保证囊括每个研究方向。②资料的典型性: 若研究成果发表年份相近, 适当考虑研究成果的被引次数和来源期刊影响力, 对于发表年份较新的研究成果则优先考虑其内容与本文研究主题的相关性。③范畴的丰富性: 若资料B的内容覆盖资料A的内容, 则剔除资料A, 若资料B与资料A均对本研究构建学术评价理论产生有益启示且不存在完全包含关系, 则将资料A和B同时纳入。该步骤迭代进行。④样本量的适切性: 本文参考在深入访谈中确定数据饱和度的方法, 经验表明深入访谈的人数应为15~25人为宜<sup>[36]</sup>, 在学术评价理论方面选择25篇中英文学术论文, 在评价实践方面选取23篇学术论文及2项报告, 合计50项研究资料进行扎根分析。

## 4 数据分析

### 4.1 开放编码

开放编码是对研究资料逐字逐句进行阅读和抽取, 从中产生初始概念并提炼出范畴, 在开放编码过程中, 需要识别研究资料中的一切理论可能, 保证足够的开放性<sup>[37]</sup>。首先由作者及另一位编码者对研究资料进行双盲编码<sup>[38]</sup>, 依照最大可能性原则提取与研究主题相关的原始语句形成代码, 所形成的代码既贴合了原始语句, 又在原始语句的基础上作出进一步的抽象化处理。形成代码后进一步将其提炼为235个概念(源于理论的概念135个, 实践100个), 详见附表1、附表2(具体请见本刊网站 <http://qxb.istic.ac.cn/CN/volumn/home.shtml>), 部分示例如表1所示。依照意思相同或相近的原则, 两位编码者对235个概念进行再次的阅读和检验, 同义词归并后整理出153个概念。从代码到概念的提炼过程体现了对代码本质属性的认知。

确定了153个概念后对其进一步归纳, 找出概念之间的内在联系, 将本质属性相同的概念归纳为同一个类别形成范畴(初始评价要素)。从概念到范畴的过程建立在对概念的深层次理解的基础上, 有助于进一步挖掘出评价活动中最重要的评价要素。开放编码形成的范畴及理论依据详见表2。

### 4.2 主轴编码

主轴编码是对开放编码提炼出的范畴进行聚类分析, 发现范畴之间的潜在关系, 根据不同范畴在层次上的包含关系对其进行归纳总结, 形成主范畴。主轴编码要解决的问题是“谁、为什么、怎样、结果如何”等问题<sup>[37]</sup>。本文根据18个范畴(初始评价要素)的内涵进行层次的梳理和类别的归纳, 最终形成了8个主范畴(关键评价要素)。

Simpson等<sup>[59]</sup>指出, 医疗环境中复杂适应性系统的特点之一是包含多样的、全面的、相互联系的行为者(agents), 包括服务的提供者、接收者、管理者、赞助者等; 行为者不断与环境互动且在环境的影响下不断发展。将此观点应用到评价领域, 评价语境作为评价活动发生的环境, 对所有评价要素产生影响。评价主体不仅与评价客体相互联系和作用, 而且充分尊重和采纳评价客体利益相关者的观点。评价主体、评价客体以及利益相关者处于与评价语境的互动过程中, 并受到评价语境的影响。评价语境、评价主体、评价客体、利益相关者4个范畴的抽象层次较高, 形成4个主范畴。

Stufflebeam<sup>[60]</sup>提出的CIPP(context-input-process-product)模型中确定了背景-投入-过程-成果的评价路径, 其中背景评价解决“需要做什么”, 明确了目标人群和需求; 投入评价考虑“应该怎么做”, 明确了最可能达到预期目标的设计和策略; 过程评价聚焦“是否正在进行”, 对评价客体进行持续的监督和反馈; 成果评价关注“是否成功实施”, 最终形成对评价客体的综合判断。Stufflebeam的CIPP模型获得了评价领域的广泛认同, 也奠定了Stufflebeam在评价领域的学术地位<sup>[61]</sup>。借鉴CIPP模型的思想, 本文从“评价目的”、“评价分类”、“评价维度”和“理论基础”4个初始评价要素中提炼出

表 1 评价理论、实践相关成果中代码到概念的形成过程示例

概念	代码	原始语句
评价语境	发生的语境	跨学科项目发生的语境极大地影响了项目过程和预期成效的质量 <sup>[39]</sup>
	语境方面的评价	关于语境方面的评价考虑财政和基础设施 <sup>[18]</sup>
	教育干预语境	在教育干预的语境中 <sup>[5]</sup>
	开放评价语境	在开放评价的语境中 <sup>[40]</sup>
评价指标	基础、特殊性评价指标	每一部分又分别包括通用性基础评价指标和高校信息公开特殊性评价指标两类 <sup>[7]</sup>
	多指标、分层次评价指标	评价指标通常应是多指标,分层次的,次序由简到繁;指标有权重之别;各指标有约束条件 <sup>[3]</sup>
	指标区间变动	还有一部分指标应该在一定的区间内变动的基础上进行评价指标体系的更科学的设计 <sup>[41]</sup>
	绩效指标	大学根据一些学术或研究绩效指标来排名 <sup>[27]</sup>
	采用的评价指标	所采用的评价指标包括:①基于论文的评价指标等 <sup>[42]</sup>
评价工具	定性及定量评价工具	定性、定量相结合,运用各种定性及定量评价工具,量化的定量分析,同行评议与描述的定性分析 <sup>[43]</sup>
	SPSS 等工具	一些图书馆还使用 SPSS、CiteSpace、HistCite、BibExcel、NVivo、NetMiner 等工具 <sup>[42]</sup>
	可信、有效的评价工具	方法论上,评价和研究需要可信、有效的工具 <sup>[44]</sup>
	评估工具	主要的评估工具是问卷调查,检查和审计 <sup>[45]</sup>
	不同评价工具	评价模型的优势在于方法的结合及不同评价工具的使用 <sup>[46]</sup>
评价对象	评价对象的特定性	基于评价对象的特定性 <sup>[47]</sup>
	不同层次评价对象	根据不同的划分标准,各个评价对象又可以划分成不同的层次 <sup>[3]</sup>
	评价项目过程的三个层面	采用此标准来评价项目过程中的三个层面,个人、组织和管理者 <sup>[39]</sup>
	评价对象的总体水平	因为单项指标并不能反映评价对象(本文主要指文献资源建设)的总体水平 <sup>[48]</sup>
评价模型	绩效评价模型	根据绩效评价目标,构建基本的绩效评价模型 <sup>[47]</sup>
	为评价提供模型	社会学习作为一个框架为跨学科研究过程的评价提供了有效的模型 <sup>[39]</sup>
	评价模型分析	通过评价模型来分析科学与技术机构的社会、环境和经济影响 <sup>[49]</sup>
	评价模型使用	出于此原因,使用的评价模型可以知晓学生、教师如何获取知识 <sup>[46]</sup>
	模型的潜在变量	在该模型中,每个潜在变量具有许多相应的测量指标 <sup>[50]</sup>
.....	.....	.....
合计数量		
153	235	235

“评价思想”,类似于CIPP模型中的背景评价,为评价活动的开展提供宏观的思想指导,在评价活动实施前确定整体的评价基调,明白“需要评价什么”;从“评价标准”、“评价制度”和“质量控制”3个初始评价要素中提炼出“评价条件”,类似于CIPP模型中的投入评价,实现对评价活动的约束和规范,规定评价活动应该如何有序开展,解决“应该如何评价”的问题;从“评价方法”和“评价指标体系”2个初始评价要素中归纳出“评价实施”,类似于CIPP模型中的过程评价,是评价具体的落实过程,包括明确的评价方法和评价指标体系;从“评价数据”和“评价结果”2个初始评价要素中提炼出“评价产出”,类似于CIPP模型中的成果评价,达成了评价目的。主轴编码形成的主范畴及内涵详见表3。

4.3 选择编码

选择编码过程是在对主范畴(关键评价要素)

进行系统分析后,选择一个核心范畴来统领所有主范畴,同时需明确关键评价要素之间的关系,并通过清晰的脉络进行整合,完成学术评价理论的建构<sup>[37,62]</sup>。本文构建学术评价理论的目的是对学术评价活动产生直接的指导作用,确保其科学有序地开展,所有主范畴(关键评价要素)都是为了清晰描述学术评价活动的整体过程。故本文确定“学术评价活动的开展”作为核心范畴,围绕这一核心范畴,借助复杂性理论的基本原理和关键概念,将评价语境、评价主体、评价客体、利益相关者、评价思想、评价条件、评价实施和评价产出8个主范畴按照合理的逻辑关系形成完整的“故事线”。

基于复杂性理论的视角,本文将整个学术评价活动视为复杂系统,评价要素是复杂系统中的构成部分。学术评价活动是一个动态发展、高度复杂的过程,借助复杂性理论中的基本原理可以用于梳理其内部各元素之间的关系<sup>[59]</sup>。复杂性理论的中心思想之一是“自组织性”,即一个系统内的各个部分



表 2 开放编码形成的范畴及理论依据

编号	范畴	概念	理论依据
1	评价语境	系统环境; 活动环境; 评价语境; 环境依赖; 现实背景; 总体环境; 语境决定; 研究环境	Klein <sup>[51]</sup> 指出由于多学科-学科内-跨学科环境的异质性导致不可能存在最佳的研究绩效评价程序
2	评价目的	评价目标; 评价导向; 基本目的; 框架用途; 示范目的; 提供引导; 结果用途; 应有作用; 评价应用; 满足需求; 评价用途; 目标权重; 政策参考; 主要目的; 参考作用; 信息提供	叶继元 <sup>[3]</sup> 认为评价目的是评价要达到的预期希望和原则要求, 是评价的龙头
3	评价分类	分类评价; 对象分类; 领域划分; 分类依据; 评价类型; 分类标准; 评价层面; 评价类别	赵伟等 <sup>[52]</sup> 认为开展人才评价时应遵循分类评价的原则, 根据各类型评价对象的不同特点和成长规律
4	评价性质	基础性、发展性评价; 细节评价; 形成性、总结性评价	Dark 等 <sup>[53]</sup> 认为评价可以划分为形成性评价和总结性评价; 张敏等 <sup>[15]</sup> 将评价划分为基础性和发展性评价
5	评价时间	评价期限; 评价时期; 评价时域; 评价时长; 时间维度; 评价周期; 评价时机	Klautzer 等 <sup>[53]</sup> 认为评价的时间不宜太早或太晚
6	评价方式	参与式评价; 认知评价	Cousins 等 <sup>[54]</sup> 提出参与式评估是基于利益相关者的模型的扩展; 郑烨等 <sup>[55]</sup> 提出从评价者的认知角度进行评价
7	评价维度	评价组合; 绩效评价维度; 维度构成	叶继元 <sup>[3]</sup> 认为评价的维度应该突破定量和定性两个维度的限制, 转向形式、内容和效用评价
8	理论基础	理论依据; 基础理论; 理论借鉴; 逻辑起点; 理论立场; 理论使用; 基于理论; 模型依据; 理论背景	Anderson 等 <sup>[46]</sup> 以复杂理论为依据构建了评价模型
9	评价主体	评价者; 评价开展者; 评价专家; 多元主体评价	文庭孝 <sup>[9]</sup> 提出不同层次和不同类型的评价主体构成了评价主体体系
10	评价客体	评价对象; 客体代理人; 参与者	叶继元 <sup>[3]</sup> 认为应该科学清晰地界定和分类评价客体
11	利益相关者	利益相关者观点; 关键利益相关者; 利益相关者认知; 利益相关者群体; 利益相关者反馈	Millstone 等 <sup>[23]</sup> 认为参与式监测和评价要求关键利益相关者的活跃参与
12	评价方法	评价算法; 方法体系; 量化考核; 访谈评价; 混合方法; 评价工具; 数据收集方法; 方法途径; 基本方法; 数理方法; 问卷评价	张士强等 <sup>[56]</sup> 提出采用定性与定量相结合、自我评价与专家组评价相结合及模糊判断的评价方法
13	评价指标体系	基本指标; 主要指标; 间接指标; 指标指数; 分层指标; 指标降维; 评价权重; 定性指标; 指标的选择、使用和解释; 指标标准化; 评价变量; 具体指标; 指标等级; 指标赋分; 指标权重; 指标测算; 基础变量; 计算指数; 测量变量	辛志勇等 <sup>[57]</sup> 认为应该确定具有操作性的评价指标体系
		逻辑模型; 评价模型; 评价框架; 评价体系; 层次结构	
14	评价标准	一般标准; 评价原则; 灵活性标准; 组织、技术标准; 构建原则	Carr 等 <sup>[39]</sup> 基于先前的研究, 形成了框架中的评价过程和结果方面的标准
15	评价制度	评价规则; 管理制度; 制度化、规范化; 规范与制度保障	邱均平等 <sup>[58]</sup> 认为我国目前还不具备有效地进行同行评议与社会公议的社会基础和制度保障
16	质量控制	管理监控; 评价质量影响因素	文庭孝 <sup>[9]</sup> 提出科学评价的质量控制要求评价机构进行质量控制
17	评价数据	要素信息; 操作化数据集合; 数据来源; 数据收集; 搜集信息; 评价数据结构; 数据分布; 数据集; 评价资料; 反馈信息; 量性、质性数据; 数据信息; 指标数据; 数据提供; 标准化数据; 数据分析	Walther 等 <sup>[40]</sup> 认为应重视评价数据结构的标准化
18	评价结果	结果展现; 结果映射; 结果产生; 结果用途; 结果分析; 结果处理; 评价结果影响因素	Bergsmann 等 <sup>[18]</sup> 认为评价的结果必须满足潜在使用者的信息需求

表3 主轴编码形成的主范畴及内涵(关键评价要素)

主范畴 (关键评价要素)	副范畴 (初始评价要素)	内涵
评价语境	评价语境	评价活动发生的环境
评价主体	评价主体	评价的实施者
评价客体	评价客体	被评价的直接对象
利益相关者	利益相关者	与评价客体存在直接利害关系
评价思想	评价目的	评价目的决定评价的方向及流程,为评价活动的开展提供全程的思想指引
	评价时间	按照评价时长和时间节点对评价进行分类,为评价活动的开展提供全程的思想指引
	评价分类	评价性质 按照评价所属性质对评价进行分类,为评价活动的开展提供全程的思想指引
	评价方式	按照评价方式对评价进行分类,为评价活动的开展提供全程的思想指引
	评价维度	评价维度包括形式、内容和效用评价,是评价思想的组成部分
评价条件	理论基础	任何科学的评价活动必须依托明确的理论基础来提供指导思想和逻辑框架,是评价思想的组成部分
	评价标准	评价标准用于指导后续评价指标的细化,对评价活动的约束和规范
	评价制度	评价制度为评价的顺利开展提供制度保障,对评价活动的约束和规范
	质量控制	通过对评价理论和方法等方面的评价达到对评价活动质量的监管,对评价活动的约束和规范
评价实施	评价方法	评价方法是指评价活动中实施的具体方法、手段和工具等,将评价落实
	评价指标体系	评价指标体系是基于细化的评价指标,遵循内在的发展规律及因果联系,形成指标的整体框架,属于评价实施过程
评价产出	评价数据	评价数据主要包括评价过程中相关信息的搜集、分析、解释和利用,属于评价产出
	评价结果	评价结果包括评价主体对评价客体所作的价值判断的结果、评价后发生的变化等,属于评价产出

根据自身的角色定位,自动地与其他部分相互作用,并形成一個稳定的运行模式,关于这个系统的运作并没有明确的规定每个组成部分应该发挥什么样的作用或如何发挥相应的作用<sup>[63]</sup>。学术评价活动的开展也应该达到自组织性的要求,让每个评价要素根据不同的定位发挥不同的作用,以特定的方式形成稳定的运行秩序,相互协作共同完成复杂的评价工作<sup>[64]</sup>。

复杂系统的“开放性”特征描述了环境因素是复杂系统中的一部分,整个系统受到环境的影响<sup>[65]</sup>。任何复杂系统在与所处环境互动的过程中受到来自环境的反馈,在不同时间、社会和地理环境中的系统必然产生差异,因此,整个学术评价活动的开展过程均受到“评价语境”的影响。复杂性理论认为一个复杂的系统是由多维度、相互作用的行为者和行为过程基于特定的关注点或功能而形成的<sup>[66]</sup>。将此观点应用于评价活动的开展,“多维度、相互作用的行为者”即评价主体、评价客体及其利益相关者,“行为过程”即评价条件、评价实施和评价产出,“基于特定的关注点或功能”即整个评价活动基于的评价思想。在评价语境的影响下,评价主体通过与评价客体及其利益相关者相互联系和作用,形成明确的评价思想,并在评价思想的指导下确定评价条件,对评价实施进行约束和规范。在评价思想的形成、评价条件的确定和评价实施的进

行过程中产生相应的评价数据,并通过对客体的评价形成评价结果。

本文构建的学术评价理论体现了复杂性理论“动态性与稳定性并存”的基本原理。学术评价理论要素中,评价思想是相对稳定的部分,通常在确定评价条件和评价实施前已经形成且不会轻易改变。评价语境、评价主体、评价客体、利益相关者一直处于相互影响和作用中,评价主体根据评价数据和评价结果的反馈,通过与评价客体及利益相关者的交互,在必要的时候对评价条件、评价实施进行调整和优化。学术评价理论模型如图2所示。

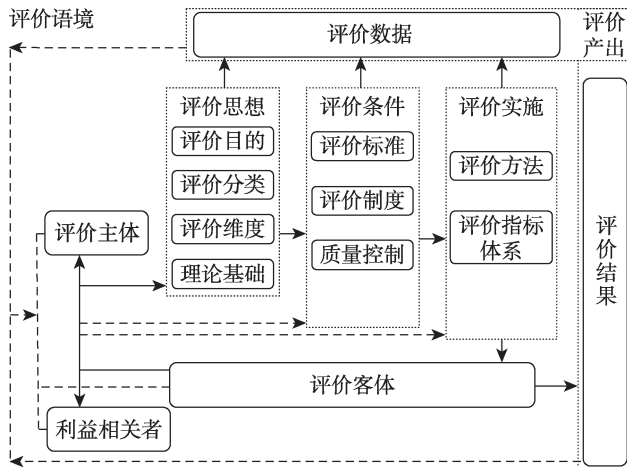


图2 学术评价理论模型



4.4 理论饱和度检验

为了保证学术评价理论模型中评价要素及相互关系的信度和效度，本文通过对 50 篇样本资料以外的学术评价理论与实践的相关研究成果重复开放

编码、主轴编码和选择编码的过程进行理论饱和度检验，均未发现形成新的评价要素及关系，表明理论模型中的范畴已经完备，理论模型是饱和的。列举 2 个代表性例子作为举证，如表 4 所示。

表 4 理论饱和度检验举例

参考文献	原始语句	代码	概念
文献[14]	采用整体分析方法识别研究生教育收益	整体分析方法	评价方法
	构建多维框架衡量研究生教育收益	多维框架	评价指标体系
	根据上述研究生教育收益评价方法并结合有关教育收益研究的成果,现对研究生教育收益加以分类	教育收益分类	评价分类
文献[28]	本评价框架是为研究的公共投资提供问责制并揭示公共投资的成效、提供基准信息并建立声誉标准、为研究提供资金的选择性分配	评价框架提供信息	评价目的
	通过实地访问,访谈评价对象	访谈评价	评价方法
	了解评价对象在提交评价材料之前的准备过程,理解他们对于收益、挑战和结果的看法	理解评价对象看法	评价对象

5 学术评价理论

5.1 评价语境

评价语境是指评价主体和客体所处的环境，是评价活动的发生地，包括社会环境、学术环境、政治环境、历史环境等<sup>[58]</sup>。评价语境会对评价客体造成影响，例如，跨学科项目发生的情境极大地影响了项目过程和预期成效的质量<sup>[39]</sup>。评价语境的复杂性为评价主体开展评价活动带来难度<sup>[47]</sup>。例如，多学科-学科内-跨学科环境的异质性导致不可能存在最佳的研究绩效评价程序，必须视具体情况而定，可见评价语境的异质性对评价活动的影响<sup>[51]</sup>，且跨学科的评价环境内科研人员及团队的构成更加复杂、科学研究的周期更长、科学研究成果的应用范围更加广泛，导致评价工作的开展需要考虑更多的因素<sup>[67]</sup>。复杂性理论中的“涌现性”认为多个要素组成的复杂系统内，复杂系统与环境在非线性相互作用和影响的过程中产生了新的功能和结构<sup>[68]</sup>。结合上述观点，本研究认为学术评价活动的整个过程均在评价语境的影响之下，各个评价要素均与评价语境产生联系。

5.2 评价主体、评价客体及利益相关者

5.2.1 评价主体

评价主体是学术评价活动的组织者、参与者和管理者<sup>[6]</sup>，需要掌握评价的逻辑和方法并对评价客体进行系统的数据收集和交流<sup>[69]</sup>。评价主体有必要保持整体上的紧密联系，按照相应领域内的秩序形

成评价主体网络，用于共享评价数据、改善评价方法、完善评价过程，避免针对同一评价客体由于信息共享的不及时等原因进行重复评价，也要防止各自为政及评价活动执行时的零碎性。评价主体的权重分配按照具体情况对多元评价主体进行合理的权重分配，例如，学术期刊的评价主体主要有高校、科研机构、数据库商、国家管理部门和专业学会等<sup>[70]</sup>，应赋予不同主体以相应的权重。另外，评价主体的性质应尽量保持纯粹，避免行政等力量的干预，制约评价活动的开展。

5.2.2 评价客体

评价客体指的是被评价的对象，涉及的范围极为广泛<sup>[3]</sup>。评价活动的科学性基于评价客体分类的合理性，因为不同评价客体具有明显不同的属性和特征。评价客体可以按照性质、特征、发展规律划分为成果、项目、学科、机构、经费、人员、管理、政策等。对于同一类评价客体，也要考虑其内部的不同细分标准，例如，对于出版物的评价应该综合考虑出版物的类型、篇幅和内容等。

5.2.3 利益相关者

评价客体往往有各类利益相关者<sup>[71]</sup>。借鉴复杂性理论中的“全面性”、“系统性”观点，评价客体的各类利益相关者需要纳入调研范围，保证对评价客体最为全面、客观和真实的评价，也有助于从侧面了解评价语境的相关信息。利益相关者的选择和参与是开展评价活动前必须确定的内容，具体包括利益相关者的参与方式及程度、识别和选择，以及在计划或设计评价活动中承担的角色及职责等<sup>[72]</sup>。

### 5.2.4 评价主体、评价客体及利益相关者的交互

评价主体需要基于客体的特征展开评价<sup>[47]</sup>。复杂性理论中的“初值敏感性”要求<sup>[73]</sup>重视起始状态,即评价主体有必要了解客体的初始状态或历史状态。与此同时,复杂性理论中的“动态性”要求避免以静止的眼光看待事物,即以动态的视角看待客体,认为客体时刻处于发展变化的过程中<sup>[74]</sup>。复杂性理论认为系统内各组成部分的相互作用会为系统带来新的属性或模式(涌现性)<sup>[63]</sup>,且由于“吸引子”的存在,整个系统呈现了长期不可预测性<sup>[63]</sup>。涌现性和吸引子提供了理解复杂系统的稳定性和变化性的框架,要求评价主体对评价客体从全局角度进行长期稳定的跟踪评价并根据评价客体的变化和反馈实时调整评价方案。评价客体作为复杂系统中的子系统,也存在复杂系统的“非周期性”特征,其变化不会体现出鲜明的周期性而呈现演变的不规则性<sup>[75]</sup>。评价活动的“非线性”特征导致评价客体的细微变化也可能对评价结果产生巨大的影响。

评价主体也需要关注客体的利益相关者<sup>[46]</sup>。复杂性理论中的“系统多样性”概念<sup>[76]</sup>描述了系统内的各个部分往往具有不同的背景、观点、经历等,这种多样性能够提高系统适应和处理复杂变化的能力。在学术评价活动中,评价主体和评价客体及利益相关者在评价过程中也必然存在背景、认知、经历等方面的差异。评价活动要尊重、调节和采纳评价主体和评价客体及利益相关者的观点,使得评价主体和评价客体及其利益相关者针对评价思想的形成、评价条件的确定和评价实施的进行形成清晰而统一的共识。

## 5.3 评价思想

### 5.3.1 评价目的

评价目的是学术评价活动开展的整体目标,决定评价的方向及流程,对评价方法的选择、评价标准和评价指标体系的制定产生直接的影响<sup>[77]</sup>。评价主体根据对评价客体的调研确定评价目的,评价目的可以划分为价值导向<sup>[27]</sup>和应用导向<sup>[49]</sup>。①价值导向:价值导向的评价没有明确需要解决的问题,重在价值做出判断,包括当前状况、最终效果及产生的影响等。例如,《美国新闻和世界报道》的全球大学排名通过对全球范围内大学的学术研究水平和整体声誉进行比较,为社会提供关于大学整体状况的认知<sup>[24]</sup>。②应用导向:应用导向的评价是通过系统化的评价,提供管理信息、决策支持、预测信

息,总结经验教训、提出改进建议等。例如,Hartmeyer等<sup>[78]</sup>探索了正规户外科学教学中的形成性评价,提供了下一步教学的决策基础。

### 5.3.2 评价分类

学术评价活动中,在实施评价前应确定评价所属的类型,从而对整体的评价活动进行明确定位,以指导后续评价的实施。确定评价分类是形成评价思想的重要组成部分,往往根据实际情况采用不同的分类标准<sup>[43]</sup>。除了根据评价客体进行分类外,还可以将评价时间、评价性质和评价方式作为分类标准。①评价时间:按照评价时间的跨度,可以将评价划分为短期、中期和长期评价,具体年限长度由各领域按照实际状况进行规定<sup>[79]</sup>。按照评价时间的节点,还可以将评价划分为事前、事中和事后评价<sup>[80]</sup>。②评价性质:评价性质分为基础性、发展性评价<sup>[15]</sup>,或形成性、总结性评价<sup>[5]</sup>。③评价方式:按照评价方式,可以将评价划分为参与式评价和认知评价,参与式评价是指利益相关者参与评价设计、数据收集、数据分析和交流<sup>[23]</sup>,认知评价是根据评价主体对评价客体的主观认知进行评价<sup>[55]</sup>。

### 5.3.3 评价维度

评价维度包括形式、内容和效用评价。形式评价是对评价客体的外部特征进行定性或定量的评价,该类评价往往可量化和可操作化的程度较高;内容评价是对评价客体的本质特征进行评价,可量化和可操作化的程度较形式评价有所降低,一般以同行评议等定性方法为主辅以定量方法,多为组织性评价流程,评价结果更加翔实可靠;效用评价是指对评价客体发挥的实际作用、效益或价值加以衡量,强调通过一定的时间和实践来评价<sup>[3]</sup>。

### 5.3.4 理论基础

任何科学的评价活动必须基于明确的理论基础,具体问题具体分析是理论基础选择的基本原则,比如,根据不同的评价目的、评价客体、评价语境等因素加以选择。理论基础可以为评价主体提供评价的认识论视角、选择评价问题和相应方法、决定以何种方式利用评价结果等<sup>[81]</sup>。理论的阐述往往是抽象的,识别和运用可以指导评价活动的特定理论需要评价主体对相关理论和评价实践有充分的了解,随之将其转换为清晰的指南和明确的概念。目前,具体描述如何将理论应用到评价实践中的工作甚少,普遍的情况是评价主体对理论基础的关注不

够,认为评价工作的开展不需要借助理论框架,或无法准确识别特定评价活动中应该遵循的理论基础<sup>[82]</sup>。

### 5.3.5 评价思想的形成

评价主体在与评价客体及利益相关者的交互中形成评价思想。评价主体首先需要借鉴已有研究,通过相关领域的文献调研、业内专家咨询等方式了解该评价领域,同时应该在充分调研评价语境<sup>[39]</sup>和了解评价客体<sup>[3]</sup>的基础上,纳入评价客体的利益相关者的观点<sup>[23]</sup>,形成评价思想。评价主体应与评价客体及利益相关者(相关部门、行业、产业界、社团及公众)充分沟通,从不同角度广泛搜集与评价客体相关的信息和观点,将评价客体的每一方利益相关者的观点纳入进来促进对评价客体的全方位考量<sup>[83]</sup>,提前对可能影响到评价客体的因素进行调研从而了解其内在关联。评价主体形成明确的评价思想后会产生相应的文本型评价数据。

## 5.4 评价条件

### 5.4.1 评价标准

评价标准是评价主体对客体的属性在质上的规定<sup>[7]</sup>,是评价主体在评价活动中应用于客体的价值尺度和界限。参照特定的评价标准可以判断评价活动是否能够达到评价目的,还可以用于指导后续评价指标的细化<sup>[3]</sup>。在某些学术评价活动中,评价主体会将已有的评价指标作为评价标准<sup>[84]</sup>,或者基于先前的研究,修订形成用于评价过程和结果的标准<sup>[39]</sup>。确定评价标准之后,评价主体还应对其操作化<sup>[39]</sup>。

### 5.4.2 评价制度

评价制度为评价的顺利开展提供保障,例如,通过相应制度来规范评价术语,利于后续评价结果的分析和利用<sup>[85]</sup>。评价过程中应该注重完善评价主体的背景审核、责任追究、随机抽取、定期轮换等制度,保证评价过程的公平。评价制度具有较强的约束力,评价过程中的相关人员需共同遵守。评价制度可以从层次上划分为总体制度、一般制度和具体制度。总体制度是关于评价体系的宏观制度,如监督和公示制度;一般制度是评价过程中需要遵循的基本制度,如回避和遴选制度;具体制度指评价过程中针对特定评价客体的规章制度<sup>[43]</sup>。评价制度不仅能够保障评价活动的正常开展,还能够促进良性评价环境的形成。

### 5.4.3 质量控制

质量控制是指对评价理论、方法、体系的评价,即对评价的评价<sup>[6]</sup>。其实现方式既包括对各个阶段连接处的监控以及时修正评价方案、技术方法、评价程序等方面可能存在的问题,也包括评价结束后对评价方案、评价流程、评价结果等方面进行分析并提出改进建议<sup>[86]</sup>。对评价的质量控制要求从以下几个方面考虑:评价主客体及利益相关者对评估的指导思想(如评价的过程和目的)是否形成共同的和全面的理解,受影响/关注的群体是否得到精准识别并有效的处理(如围绕评价的意图和结果征求利益相关者的意见并使其保持知情的方式),评价中各方权利是否得到合同和法律方面的保障(如有什么保障措施确保评价服务于所有层次的利益相关者)、是否有详细和灵活的评价方案(如评价主体采用何种数据收集工具和技术来确保处理关键的评价问题)、是否有完善的管理方案(如实施何种监督和控制来确保评价主体投入时间、精力和声誉开展评价),评价是否道德方面服务于具有社会价值的目标(如评价主体确保公平地解决和尊重所有利益相关者的需要和权利,并适当考虑其性别、种族和语言背景的措施),评价主体是否对评价结果有建设性使用(如是否有适当的规定来根据不同受众的需要调整报告)。

### 5.4.4 评价条件的确定

评价思想是在评价主体调研评价语境、了解评价客体的基础上形成的,指引评价标准、评价制度以及质量控制等评价条件的确定。依据不同的评价目的会形成不同的评价标准<sup>[3]</sup>,并进而影响评价制度以及质量控制策略与方法。评价分类可以对整体的学术评价活动进行明确定位<sup>[43]</sup>,评价维度指定了评价需要考虑的角度<sup>[88]</sup>,理论基础为评价条件的制定提供评价的方法论依据。总之,评价目的、评价分类、评价维度、理论基础以及评价语境的协同作用确定了评价条件。

## 5.5 评价实施

### 5.5.1 评价方法

评价方法是评价实施过程中采用的具体手段、工具等<sup>[44]</sup>,可分为定性和定量两大类。评价主体可以根据不同的评价目的选择特定的评价类型,而对于不同类型的评价,往往需定位到适切的评价方法<sup>[47]</sup>。鉴于实际评价活动中通常需根据具体语境做



出灵活调整,因此评价目的、评价分类和评价方法之间不是严格的线性指向关系。除了常用的定量评价方法之外,评价主体还可以通过问卷、访谈<sup>[18]</sup>以及焦点小组等定性评价方法搜集评价客体面临的挑战,从而深入揭示评价客体目前存在的问题或有待提高的部分,以加深对评价客体的认识。评价实践中,宜通过定性和定量评价方法的结合运用以提高评价的信度和效度。

### 5.5.2 评价指标体系

基于复杂性理论,系统功能并不是系统内各组成部分的简单相加,因此对于评价指标的制定需要从宏观、中观及微观进行通盘考虑,并依照合理的评价框架形成评价指标体系<sup>[89]</sup>。①评价指标:评价指标在内容方面要考虑因素间的关联程度及指标值规范化。当评价目的难以准确量化时,指标的细分也将面临挑战,这时应该着重理解“模糊”的概念<sup>[90]</sup>,可以通过意义建构变得易于控制和管理。例如,泰晤士高等教育2018世界大学排名报告中的研究收入指标受到国家政策和经济环境的影响,存在一定的“模糊性”<sup>[25]</sup>。评价指标的形式应该达到简洁、明晰和多层次的要求。此外,评价指标必须具备可操作化。②评价指标体系:评价指标体系的形成应遵循内在的发展规律及因果联系<sup>[91]</sup>,应具有环境适应性、系统性、针对性、可操作性等,应能够客观反映评价客体发展状况,并预测评价客体发展趋势<sup>[92]</sup>。

### 5.5.3 评价实施的约束与规范

不同评价方法的应用前提和适用范围有显著差异<sup>[20]</sup>,评价方法的使用必须遵循一定的程序<sup>[93]</sup>,尤其是在评价方法的选择存在争议时<sup>[94]</sup>。评价主体应参照严格的标准,根据实际情况制定出适用的评价框架<sup>[53]</sup>,评价框架的效度可以为评价指标体系的有效性提供保障<sup>[39]</sup>,进而针对评价客体产生可信的评价结果。评价条件中的评价标准、评价制度和质量控制可以对评价实施进行约束和规范,尽量避免评价实施过程中可能出现的评价方法选择等方面的问题<sup>[86]</sup>。

### 5.5.4 评价结果的产生

评价实施作用于评价客体产生评价结果。在评价实施过程中,评价主体需综合采用多种评价方法全方位收集评价客体的量化数据与质性数据<sup>[57]</sup>,判断评价数据的完整性、真实性以及可靠性等<sup>[27]</sup>,保

证数据结构的标准化<sup>[40]</sup>,并采用科学的方法进行数据分析<sup>[45]</sup>。根据多角度、多元化的评价指标体系对评价客体进行全面考察<sup>[95]</sup>,力求产生有根据的、可操作的、可信的和有用的评价结果<sup>[5]</sup>。

## 5.6 评价产出

### 5.6.1 评价数据

评价数据是指整个评价活动中所产生的各类定量和定性数据,是实时的评价产出<sup>[55]</sup>,贯穿于评价思想的形成、评价条件的确定和评价实施等过程。评价主体根据对评价客体的调研形成明确的评价思想用于指导后续评价的实施,此阶段形成评价目的、评价分类、评价维度和理论基础相关的数据,以无结构的文本为主。评价思想形成之后,为了确保评价实施的科学性与合理性,通过确立完整的评价标准、评价制度和质量控制达到对评价实施的约束,此阶段也将形成规范性质的评价数据,同样以无结构的文本为主。评价实施中评价数据产生于评价方法的选择和评价指标体系的制定过程,更多地产生于面向评价客体进行各方面信息的搜集、分析与解释等过程,涵盖无结构的文本以及有结构的量化数据。

### 5.6.2 评价结果

评价结果是最终的评价产出,包括对其的分析和利用<sup>[18,95]</sup>。对评价结果的分析需要考虑评价结果产生的原因和所在的语境,挖掘出其中蕴含的政治、经济、文化等方面的意义,力求让一次评价活动能够产生长远的影响<sup>[46]</sup>。评价结果应该具有实时性、共享性、动态性、公开性的特征,保证评价结果可供查询、申诉、审核、追溯和监督,避免围绕同一评价客体进行多头评价的现象,同时最大限度地避免功利化,杜绝错用和滥用。

### 5.6.3 评价数据和评价结果的反馈

复杂性理论中的“自适应系统”是在系统运行过程中,不断采集控制过程信息,确定被控对象的当前实际工作状态,优化性能准则,产生自适应控制规律,从而实时地调整控制器结构或参数,使得系统始终以最优的方式运行,按照环境的变化调整自身以发挥最佳功能<sup>[96]</sup>。将此概念映射到评价过程中,评价主体、客体以及利益相关者接受来自评价数据和评价结果的反馈,若发现评价过程中存在问题,则进一步通过三者的交互,对评价条件和评价

实施进行反思和改进。例如,评价主体应用特定评价方法搜集评价客体的数据后,发现评价条件或评价实施有调整之必要,使得整个评价过程成为一个循环系统<sup>[46]</sup>。评价实践中也会出现部分极端的情况,比如,上述反馈过程引发对评价思想的反思,使得先前的评价工作难以为继。

#### 5.6.4 评价结果的利用

评价结果产生之后,评价主体在宏观层面审视评价过程的同时,也不能忽视低层次的微观现象,即评价结果不仅仅要检验评价客体的真实状态,更要挖掘背后的原因,总结相关经验,用于评价客体的改善。为了有效利用评价结果来指导实践或提供决策支持<sup>[78]</sup>,评价结果必须以评价客体及利益相关者易于理解和使用的描述<sup>[18]</sup>。例如,将评价结果与评价客体整体的发展规划和长远的战略目标相结合,以利于评价客体的良性发展<sup>[97]</sup>。

## 6 结 论

本文采用扎根理论的研究方法,通过开放编码从学术评价理论与实践相关成果中发掘235条代码,并将其抽象化处理为235个概念,并进一步提炼为153个概念。基于概念之间的内在关系,将本质属性相同的概念归纳为18个范畴(初始评价要素)。在主轴编码过程中,通过分析范畴之间的潜在关系,根据范畴的内涵进行层次梳理和类别归纳,形成了8个主范畴(关键评价要素)。在选择编码阶段,本文以“学术评价活动的开展”作为统领所有主范畴的核心范畴,借助复杂性理论的基本原理和关键要素,将评价语境、评价主体、评价客体、利益相关者、评价思想、评价条件、评价实施和评价产出8个主范畴按照合理的逻辑关系形成完整的“故事线”,构建了学术评价理论。

扎根分析显示,在评价语境的影响下,评价主体在与评价客体及其利益相关者相互作用中形成明确的评价思想(评价目的、评价分类、评价维度、理论基础),并进一步确定评价条件(评价标准、评价制度、质量控制),对评价实施进行约束和规范。形成评价思想、确定评价条件和进行评价实施的过程中产生实时评价产出(评价数据)和最终评价产出(评价结果)。本文在厘清概念内涵的基础上规范了学术评价术语,便于学术评价领域内的无障碍沟通,并将学术评价活动中的关键评价要素以合理的方式进行组织和整合,探索了其中的逻辑关

系和内在机理,为学术评价活动的开展提供了全面的指导框架。

目前学术界对基于小样本扎根分析得到的研究结果的信度、效度以及通用性等尚存争议<sup>[98]</sup>。因此,本文构建的学术评价理论有待实践的进一步检验。

## 参 考 文 献

- [1] 刘志明. 以发展着的理论指导新的实践[EB/OL]. (2017-11-13) [2018-10-23]. <http://theory.people.com.cn/n1/2017/1113/c40531-29642265.html>.
- [2] 实践没有止境理论创新也没有止境[EB/OL]. (2018-03-07) [2019-02-20]. <http://theory.people.com.cn/n1/2018/0307/c40531-29852958.html>.
- [3] 叶继元. 人文社会科学评价体系探讨[J]. 南京大学学报(哲学·人文科学·社会科学), 2010, 47(1): 97-110, 160.
- [4] 辛涛, 李雪燕. 教育评价理论与实践的新进展[J]. 清华大学教育研究, 2005, 26(6): 38-43.
- [5] Dark M, Mirkovic J. Evaluation theory and practice applied to cybersecurity education[J]. IEEE Security & Privacy, 2015, 13(2): 75-80.
- [6] 文庭孝. 科学评价理论体系的构建研究[J]. 重庆大学学报(社会科学版), 2008, 14(3): 63-69.
- [7] 吕红. 高校信息公开政策评价理论体系的构建研究[J]. 图书情报工作, 2012, 56(18): 25-28, 92.
- [8] Dewey J. Valuation and experimental knowledge[J]. The Philosophical Review, 1922, 31(4): 325-351.
- [9] Shadish W R. Evaluation theory is who we are[J]. The American Journal of Evaluation, 1998, 19(1): 1-19.
- [10] Donaldson S I, Lipsey M W. Roles for theory in contemporary evaluation practice: Developing practical knowledge[M]// The SAGE Handbook of Evaluation. Thousand Oaks: SAGE Publications, 2006: 57-75.
- [11] Hansen M, Alkin M C, Wallace T L. Depicting the logic of three evaluation theories[J]. Evaluation and Program Planning, 2013, 38: 34-43.
- [12] Scriven M. Evaluation theory and metatheory[M]// International Handbook of Educational Evaluation. Dordrecht: Springer, 2003, 9: 15-30.
- [13] Alkin M C, Christie C A. An evaluation theory tree[M]// Evaluation Roots. Thousand Oaks: SAGE Publications, 2004: 13-65.
- [14] 张志昂. 研究生教育收益评价理论初探[J]. 教育与经济, 2003 (2): 45-48.
- [15] 张敏, 唐任伍. 我国高等教育资源利用效率评价理论研究[J]. 清华大学教育研究, 2006, 27(5): 24-28, 41.

- [16] 王俭. 当前我国教育评价理论研究存在的问题与实践误区的价值取向分析[J]. 教师教育研究, 2008, 20(6): 49-55.
- [17] 李凌艳, 李勉. 从西方教育评价理论发展的视角看我国学校评估研究[J]. 教育理论与实践, 2010, 30(2): 25-29.
- [18] Bergsmann E, Schultes M T, Winter P, et al. Evaluation of competence-based teaching in higher education: From theory to practice [J]. Evaluation and Program Planning, 2015, 52: 1-9.
- [19] 王媛, 马小燕. 基于模糊理论与神经网络的人才评价方法[J]. 佳木斯大学学报(自然科学版), 2006, 24(3): 408-410.
- [20] 赵伟, 林芬芬, 彭洁, 等. 创新型科技人才评价理论模型的构建[J]. 科技管理研究, 2012, 32(24): 131-135.
- [21] 许海云, 李长玲. 基于未确知理论与信息熵方法的技术类知识产品的价值评价体系[J]. 图书情报工作, 2009, 53(16): 65-68.
- [22] 乔欢. 数字图书馆评价理论框架构建者萨拉赛维奇[J]. 国家图书馆学刊, 2006, 15(1): 74-76.
- [23] Millstone E, van Zwanenberg P, Marshall F. Monitoring and evaluating agricultural science and technology projects: Theories, practices and problems[J]. IDS Bulletin, 2010, 41(6): 75-87.
- [24] U.S. News. How U.S. news calculated the best global universities rankings[EB/OL]. [2018-11-02]. <https://www.usnews.com/education/best-global-universities/articles/methodology>.
- [25] Times Higher Education. World university rankings 2018 methodology[EB/OL]. [2018-08-21]. <https://www.timeshighereducation.com/world-university-rankings/methodology-world-university-rankings-2018>.
- [26] Quacquarelli Symonds. The QS world university rankings[EB/OL]. [2019-03-12]. <https://www.topuniversities.com/qs-world-university-rankings/methodology>.
- [27] 上海软科教育信息咨询有限公司. Academic ranking of world universities 2018[EB/OL]. [2018-10-22]. <http://www.shanghai-ranking.com/ARWU-Methodology-2018.html>.
- [28] High education funding council for England. Evaluation of REF 2014[EB/OL]. [2018-08-22]. <https://www.ref.ac.uk/2014/media/ref/content/pub/assessmentframeworkandguidanceconsuissions/GOS%20including%20addendum.pdf>.
- [29] 南京大学中国社科院研究评价中心. 中文社会科学引文索引来源期刊/集刊遴选实施方案(试行)[EB/OL]. [2018-10-23]. <http://cssrac.nju.edu.cn/a/zlxz/20171221/2857.html>.
- [30] 上海社会科学院智库研究中心. 2017中国智库报告[EB/OL]. (2018-02-07) [2018-09-23]. <http://www.199it.com/archives/690079.html>.
- [31] Walton M. Applying complexity theory: A review to inform evaluation design[J]. Evaluation and Program Planning, 2014, 45: 119-126.
- [32] Jorm C, Roberts C. Using complexity theory to guide medical school evaluations[J]. Academic Medicine, 2017, 93(3): 399-405.
- [33] Eoyang G, Oakden J. Adaptive evaluation: A synergy between complexity theory and evaluation practice[J]. Emergence-Complexity & Organization, 2016, 18: 3-4.
- [34] Walton M. Expert views on applying complexity theory in evaluation: Opportunities and barriers[J]. Evaluation, 2016, 22(4): 410-423.
- [35] Smith N L. Improving evaluation theory through the empirical study of evaluation practice[J]. Evaluation Practice, 1993, 14(3): 237-242.
- [36] 靳代平, 王新新, 姚鹏. 品牌粉丝因何而狂热?——基于内部人视角的扎根研究[J]. 管理世界, 2016(9): 102-119.
- [37] Corbin J, Strauss A. Basics of qualitative research: Techniques and procedures for developing grounded theory[M]. Thousand Oaks: SAGE Publications, 2008: 159-160, 198-199.
- [38] 许庆瑞, 吴志岩, 陈力田. 转型经济中企业自主创新能力演化路径及驱动因素分析——海尔集团1984~2013年的纵向案例研究[J]. 管理世界, 2013(4): 121-134, 188.
- [39] Carr G, Loucks D P, Blöschl G. Gaining insight into interdisciplinary research and education programmes: A framework for evaluation[J]. Research Policy, 2018, 47(1): 35-48.
- [40] Walther A, van den Bosch J J F. FOSE: A framework for open science evaluation[J]. Frontiers in Computational Neuroscience, 2012, 6: 32.
- [41] 谭和平. 我国公益类科研院所运行评价体系——基于管理熵理论的研究[J]. 社会科学研究, 2008(5): 192-194.
- [42] 王颀. “985”高校图书馆科研评价服务实践调研与分析[J]. 图书情报工作, 2016, 60(1): 26-31.
- [43] 夏云霞, 徐涛, 翟康, 等. 研究所科研团队绩效评价的探索与实践[J]. 科研管理, 2017, 38(S1): 510-514.
- [44] Granit-Dgani D, Kaplan A, Flum H. Theory-based assessment in environmental education: A tool for formative evaluation[J]. Environmental Education Research, 2017, 23(2): 269-299.
- [45] Fjortoft N F, Schwartz A H. Evaluation of a pharmacy continuing education program: Long-term learning outcomes and changes in practice behaviors[J]. The American Journal of Pharmaceutical Education, 2003, 67(2): 35.
- [46] Anderson E, Smith R, Hammick M. Evaluating an interprofessional education curriculum: A theory-informed approach[J]. Medical Teacher, 2016, 38(4): 385-394.
- [47] 林梦泉, 陈燕, 任超, 等. 约束条件下的学科建设绩效评价理论体系探究[J]. 中国高教研究, 2018(7): 17-21.
- [48] 吉汉强, 李丽舒, 黄超云, 等. 文献资源建设绩效评价指标体系构建的实践研究[J]. 图书馆建设, 2011(4): 26-30.
- [49] de Lima R G, Lins H N, Pfitscher E D, et al. A sustainability evaluation framework for Science and Technology Institutes: An inter-



- national comparative analysis[J]. *Journal of Cleaner Production*, 2016, 125: 145-158.
- [50] Yu L P, Chen Y Q, Pan Y T, et al. Research on the evaluation of academic journals based on structural equation modeling[J]. *Journal of Informetrics*, 2009, 3(4): 304-311.
- [51] Klein J T. Evaluation of interdisciplinary and transdisciplinary research: A literature review[J]. *American Journal of Preventive Medicine*, 2008, 35(Suppl 2): S116-S123.
- [52] 赵伟, 包献华, 屈宝强, 等. 创新型科技人才分类评价指标体系构建[J]. *科技进步与对策*, 2013, 30(16): 113-117.
- [53] Klautzer L, Hanney S, Nason E, et al. Assessing policy and practice impacts of social science research: The application of the Payback Framework to assess the Future of Work programme[J]. *Research Evaluation*, 2011, 20(3): 201-209.
- [54] Cousins J B, Earl L M. The case for participatory evaluation[J]. *Educational Evaluation and Policy Analysis*, 1992, 14(4): 397-418.
- [55] 郑烨, 胡春萍. 基金委是透明的吗?——基于循证评估的科学基金信息公开实践评价[J]. *情报杂志*, 2017, 36(9): 78-85.
- [56] 张士强, 马光成. 毕业设计质量评价的研究与实践[J]. *中国高教研究*, 2001(10): 80-81.
- [57] 辛志勇, 杜晓鹏, 许晓晖. 澳大利亚学校价值观教育实效性评价实践[J]. *比较教育研究*, 2016, 38(9): 7-13.
- [58] 邱均平, 任全娥. 国内外人文社会科学科研成果评价比较研究[J]. *国外社会科学*, 2007(3): 58-66.
- [59] Simpson K M, Porter K, McConnell E S, et al. Tool for evaluating research implementation challenges: A sense-making protocol for addressing implementation challenges in complex research settings[J]. *Implementation Science*, 2013, 8: 2.
- [60] Stufflebeam D L. The relevance of the CIPP evaluation model for educational accountability[J]. *Journal of Research and Development in Education*, 1971, 5(1): 19-25.
- [61] 肖远军. CIPP教育评价模式探析[J]. *教育科学*, 2003, 19(3): 42-45.
- [62] 王建明, 王俊豪. 公众低碳消费模式的影响因素模型与政府管制政策——基于扎根理论的一个探索性研究[J]. *管理世界*, 2011(4): 58-68.
- [63] 吴今培, 李雪岩, 赵云. 复杂性之美[M]. 北京: 北京交通大学出版社, 2017: 27, 53, 104.
- [64] Trenholm S, Ferlie E. Using complexity theory to analyse the organisational response to resurgent tuberculosis across London[J]. *Social Science & Medicine*, 2013, 93: 229-237.
- [65] 苗东升. 复杂性科学研究[M]. 北京: 中国书籍出版社, 2013: 124.
- [66] Gare A. Systems theory and complexity: Introduction[J]. *Democracy & Nature*, 2000, 6(3): 327-339.
- [67] 魏巍, 刘仲林. 国外跨学科评价理论新进展[J]. *科学学与科学技术管理*, 2011, 32(4): 20-25.
- [68] 陈一壮. 试论复杂性理论的精髓[J]. *哲学研究*, 2005(6): 108-114.
- [69] Dillman L M. Comparing evaluation activities across multiple theories of practice[J]. *Evaluation and Program Planning*, 2013, 38: 53-60.
- [70] 臧莉娟. 期刊评价结果权威性影响因素分析——以人文社会科学期刊评价为中心[J]. *中国出版*, 2015(16): 11-15.
- [71] Ladyshevsky R K, Taplin R. Evaluation of curriculum and student learning needs using 360 degree assessment[J]. *Assessment & Evaluation in Higher Education*, 2015, 40(5): 698-711.
- [72] Mark M M. Building a better evidence base for evaluation theory [M]// *Fundamental Issues in Evaluation*. New York: The Guilford Press, 2008: 111-134.
- [73] 刘晓君, 李险峰, 张建刚. 一个新自治混沌系统的混沌同步控制[J]. *复杂系统与复杂性科学*, 2007, 4(4): 51-57.
- [74] 金吾伦, 郭元林. 复杂性科学及其演变[J]. *复杂系统与复杂性科学*, 2004, 1(1): 1-5.
- [75] Marchal B, van Belle S, de Brouwere V, et al. Studying complex interventions: Reflections from the FEMHealth project on evaluating fee exemption policies in West Africa and Morocco[J]. *BMC Health Services Research*, 2013, 13: 469.
- [76] 吴彤. 复杂的实在[J]. *自然辩证法研究*, 2005, 21(6): 1-4, 10.
- [77] 叶继元. 图书馆学期刊质量“全评价”探讨及启示[J]. *中国图书馆学报*, 2013, 39(4): 83-92.
- [78] Hartmeyer R, Stevenson M P, Bentsen P. Evaluating design-based formative assessment practices in outdoor science teaching[J]. *Educational Research*, 2016, 58(4): 420-441.
- [79] 严明清. 地方社科院成果评价体系及科研管理创新与思考[J]. *社会科学管理与评论*, 2009(4): 74-79, 112.
- [80] 胡琳, 刘倩, 舒予. 图书馆员国际交流项目评价体系研究[J]. *大学图书馆学报*, 2017, 35(3): 18-24.
- [81] Bonaccorsi A, Piccaluga A. A theoretical framework for the evaluation of university-industry relationships[J]. *R & D Management*, 1994, 24(3): 229-247.
- [82] Alkin M C. Evaluation theory and practice: Insights and new directions[J]. *New Directions for Evaluation*, 2003, 2003(97): 81-90.
- [83] Rallis S F, Rossman G B. Dialogue for learning: Evaluator as critical friend[J]. *New Directions for Evaluation*, 2000, 2000(86): 81-92.
- [84] 于永达, 贺贵才. 自主知识产权产品评价: 理论与实例研究[J]. *科学学研究*, 2008, 26(6): 1188-1192.
- [85] 季靖. 高校教师教学质量评价实践及启示[J]. *中国流通经济*, 2001, 15(S1): 32-35.
- [86] The Evaluation Center at Western Michigan University. Evaluation checklists[EB/OL]. [2018-09-08]. <https://www.wmich.edu/>

- evaluation/checklists,2018-03-02/.
- [87] Stufflebeam D L. Evaluation plans and operations checklist[EB/OL]. [2020-07-24]. <https://wmich.edu/sites/default/files/attachments/u350/2018/eval-plans-operations-stufflebeam.pdf>.
- [88] 李月琳, 梁娜, 齐雪. 从交互维度到交互功能: 构建数字图书馆交互评估理论模型[J]. 中国图书馆学报, 2016, 42(1): 66-82.
- [89] Phillips-Wren G, Mora M, Forgionne G A, et al. An integrative evaluation framework for intelligent decision support systems[J]. European Journal of Operational Research, 2009, 195(3): 642-652.
- [90] Dahler-Larsen P. Theory-based evaluation meets ambiguity: The role of Janus variables[J]. American Journal of Evaluation, 2018, 39(1): 6-23.
- [91] Tackett S, Grant J, Mmari K. Designing an evaluation framework for WFME basic standards for medical education[J]. Medical Teacher, 2016, 38(3): 291-296.
- [92] 吕潭华. 可持续发展评价指标体系研究[D]. 厦门: 厦门大学, 2002.
- [93] 郎永杰, 杨青, 郭芳芳. 产学研合作绩效评价: 内涵、缘起与实践[J]. 教育理论与实践, 2017, 37(24): 6-8.
- [94] Xu H, Xing L N, Huang L. Evaluation of scientific publications with hesitant fuzzy uncertain linguistic and semantic information [J]. Journal of Intelligent and Fuzzy Systems, 2015, 29(6): 2737-2742.
- [95] 陶军. 权利保障维度的国家科研资助评价体系构建——以国家社科基金为例[J]. 宏观质量研究, 2017, 5(2): 82-97.
- [96] 金福, 陈伟. 遗传算法之父——霍兰及其科学工作[J]. 自然辩证法通讯, 2007, 29(2): 86-93, 112.
- [97] Gangopadhyay P. Making evaluation meaningful to all education stakeholders[R]. The Evaluation Center, Western Michigan University, 2002: 1-3.
- [98] 陈向明. 社会科学中的定性研究方法[J]. 中国社会科学, 1996 (6): 93-102.

(责任编辑 魏瑞斌)

附表 1 评价理论相关成果中代码到概念的形成过程

参考文献	原始语句	代码	概念
林梦泉 <sup>[47]</sup>	基于评价对象的特定性	评价对象的特定性	评价对象
	本研究基于“投入-产出”理论中教育绩效评价的特殊性	基于“投入-产出”理论	理论依据
	设定评价期限,在此期间的投入相对有规律性和稳定性	设定评价期限	评价期限
	绩效评价目标设定主要基于管理者的目的	绩效评价目标	评价目标
	根据绩效评价目标,构建基本的绩效评价模型	绩效评价模型	评价模型
	作为绩效评价的基本导向	评价的基本导向	评价导向
	基于目标的体系构建应明确 3 个规则	明确 3 个规则	评价规则
	要素信息获取。绩效评价体系确定后,信息获取是关键	要素信息获取	要素信息
	根据投入产出要素的性质、绩效评价目标,需选择不同算法形成绩效评价结果	不同算法	评价算法
	绩效评价结果呈现的一般是一种比率、相对值的含义	绩效评价结果	评价结果
	以期对不同情况采用不同绩效评价方法提供参考	绩效评价方法	评价方法
	由于教育系统的复杂性,认为研究理想的绩效评价模型存在一定难度	教育系统复杂性	系统环境
吕红 <sup>[7]</sup>	在评价主体的选择上要本着利益相关、多元化、独立性、专业化、公开公平、代表性等主体性原则	评价主体的选择	评价主体
	高校信息公开政策评价客体是评价的主要内容及特定对象	高校信息公开政策评价客体	评价客体
	高校信息公开政策评价标准是指在评价高校信息公开政策之前,政策评价主体对客体的属性在质上的规定	高校信息公开政策评价标准	评价标准
	有学者依据信息政策方案评价指标体系构建原则	评价指标体系构建原则	评价原则
	指出其具有通用性的基本指标有:目标明确性、时效性、创新性、科学性、规范性、完整性、权威性、可认知程度、利益调控程度及信息政策方案可执行程度	通用性的基本指标	基本指标
	指出政策效果评价的一般标准有:效能标准、效率标准、充分标准、公正标准、适当标准	政策效果评价的一般标准	一般标准
	每一部分又分别包括通用性基础评价指标和高校信息公开特殊性评价指标两类	基础、特殊性评价指标	评价指标
	官方教育组织机构对整个评价过程进行管理、监控	评价过程管理、监控	管理监控
叶继元 <sup>[3]</sup>	其评价实施期由科研立项的第三方科研机构进行,也包括三个阶段,分别是评价前期、评价中期、评价后期	评价前期、中期、后期	评价时期
	评价主体是多元的,可以是学者、专家、同行个人或群体、学术机构、学术期刊、报纸、网站、电视、科研管理部门等	多元评价主体	评价主体
	根据不同的划分标准,各个评价对象又可以划分成不同的层次	不同层次评价对象	评价对象
	科学清晰地界定和分类评价客体,对于人文社科的“按特点分类评价”的原则具有重要意义	按特点分类评价	分类评价
	不同的评价目的决定着不同的评价标准和指标	不同评价目的	评价目的
	这些方法又被概括成定量评价法、定性评价法和定性定量综合评价法	定量、定性评价方法	评价方法
	评价指标通常应是多指标,分层次的,次序由简到繁;指标有权重之别;各指标有约束条件	多指标、分层次评价指标	评价指标
	评价标准是指人们在评价活动中应用于对象的价值尺度和界限	评价标准是价值尺度和界限	评价标准
	评价制度是有关部门制定的保证评价活动进行、要求有关人员共同遵守的规程	评价制度是共同遵守的规程	评价制度
	系统评价理论则为构建人文社科评价体系提供了直接依据	系统评价理论提供依据	理论依据
文庭孝 <sup>[6]</sup>	结合国内外人文社科评价实践及其存在的问题提出了三个新概念组合,即形式评价、内容评价和效用评价	形式、内容、效用的评价组合	评价组合
	科学评价活动以整个科研活动系统为评价对象	评价对象	评价对象
	评价制度是约束和规范科学评价活动的各种行为规则、准则的集合	评价制度是行为规则和准则的集合	评价制度



续表

参考文献	原始语句	代码	概念
	科学评价的质量控制要求评价机构建立内部的质量控制系统和外部的质量监督机制	质量控制要求内部、外部监督	质量控制
	科学评价主体是指实施评价的各类组织、机构和人员,是科学评价活动的组织者、参与者和管理者	科学评价主体是组织者、参与者和管理者	评价主体
	科学评价基础理论,主要包括科学评价的概念体系、对象体系、内容体系、理论基础和基本理论等	科学评价基础理论	基础理论
	科学评价方法体系,主要包括评价方法论、评价方法、评价模型和评价程序等	科学评价方法体系	评价方法
	科研活动环境包括外部环境和内部环境	科研活动环境	活动环境
李月琳等 <sup>[88]</sup>	识别影响用户评估数字图书馆交互功能的主要指标	识别主要指标	主要指标
	构建有效的数字图书馆交互功能评估模型	构建评估模型	评估模型
	细化李月琳等提出的三维模型中的各交互维度,明确子维度	各交互维度、子维度	评价维度
郭振宇 <sup>①</sup>	可以用模糊数学对高等教育的收益作出综合评价	用模糊数学作出综合评价	评价方法
	根据上述高等教育收益评价方法并结合有关教育收益的定义,现对高等教育收益加以分类	高等教育收益分类	对象分类
	笔者试图以劳动生产率、社会犯罪率、公民参与政治投票率等指标来评价高等教育的社会间接收益	以指标评价间接收益	间接指标
于永达等 <sup>[84]</sup>	自主知识产权产品评价的基本目的在于衡量产品的自主知识产权化程度	评价的基本目的	基本目的
	在综合计算之前,需要对评价指标指数进行归一化处理和无量纲化处理	处理评价指标指数	指标指数
	在当前国家经济统计体系基本指标基础上,构建的派生指标,作为度量自主知识产权产品的评价标准	度量产品的评价标准	评价标准
袁安府等 <sup>②</sup>	为了清晰而便于评价,应该按某些原则合理地将评价指标分为目标层、准则层与指标层等若干层次	目标层、准则层与指标层等评价指标	分层指标
	归纳和提炼出包括三个评价维度(哲学社会科学事业发展投入指标、哲学社会科学事业发展产出指标、哲学社会科学事业发展基础指标)	投入、产出、发展基础三个评价维度	评价维度
	把复杂的评价指标体系条理化、层次化,建立一个有层次的结构模型	层次的评价指标体系结构模型	评价模型
	10个评价领域和40个评价指标相对于评价目标的权重系数	评价目标的权重系数	目标权重
	归纳和提炼出10个评价领域(科研经费投入、国家社科立项课题等)	归纳和提炼评价领域	领域划分
	评价指标的理论遴选应该遵守系统性、可操作性、有效性、可比性和动态性五项基本原则	评价指标的理论遴选基本原则	评价原则
谭和平 <sup>[41]</sup>	借鉴管理熵理论的基本定义,可以建立起如下的评价模型	借鉴理论的基本定义	理论借鉴
	在评价的时域上,注重对现代科研院所发展潜力的评价	评价时域上的发展潜力	评价时域
	如何在对绩效和行为指标进行分类(部分指标的特点是越大越好,部分指标的特点是越小越好,还有一部分指标应该在一定的区间内变动才合理)的基础上进行评价指标体系的更科学的设计	指标区间变动	评价指标
张敏 <sup>[15]</sup>	本文以高等教育的功能和办学目标为依据进行分类	以功能和目标为依据分类	分类依据
	前者用于对高等教育有效资源利用效率的普遍意义上的基础性评价;后者用于对高等教育有效资源利用效率的更高层次上的发展性评价	普遍的基础性、高层次的发展性评价	基础性、发展性评价
	综合评价的方法有很多,其中应用较多的是加权评分法、专家评分法、模糊综合评价法等	综合评价方法	评价方法
	本文采用主成分分析的方法,它是一种可以对原指标进行降维的多元统计分析方法	原指标降维	指标降维
赵伟等 <sup>[20]</sup>	建立影响我国高等教育资源利用效率的层次结构体系	层次结构体系	层次结构
	依托胜任力模型理论与个体创新行为理论	依托模型、行为理论	理论依据
	不同评价方法的应用前提和适用范围尚不清晰	评价方法应用前提和使用范围	评价方法
	建立起多角度、多元化、由一系列特定组合并相互独立的指标组成的综合评价指标体系	多角度、多元化综合评价指标体系	评价指标体系

续表

参考文献	原始语句	代码	概念
严明清 <sup>[79]</sup>	开始将激励重心转向决策咨询类成果,在科研成果评价中相对提高这类成果的评价权重	提高决策咨询类成果的评价权重	评价权重
	实行对部分科研人员免除年度科研考核,拉长考核期	拉长考核期	评价时长
	改变过去的简单的数量考核,开始引入量化积分的考核方式	量化积分考核方式	量化考核
陶军 <sup>[95]</sup>	从权利视角对国家科研资助的内涵、性质和根本价值进行分析,奠定指标体系设计的逻辑起点	从权利视角奠定逻辑起点	逻辑起点
	需要在对评价对象的进程进行分析	评价对象进程	评价对象
	从时间维度识别出国家科研资助的不同阶段	从时间维度识别	时间维度
	在法治语境下,主要是指宪法和法律对国家的积极保障义务	法治语境	评价语境
	定性评价较适用于对特定对象的深入考察,一些指标宜分析描述而不能简单量化	分析描述定性评价指标	定性指标
	全国哲学社会科学规划办公室制定和颁布了一系列管理制度	制定和颁布管理制度	管理制度
	评估结果一般以文字报告展现,是对客观现象的法治价值评判	评价结果以文字报告展现	结果展现
Carr 等 <sup>[39]</sup>	跨学科项目发生的语境极大地影响了项目过程和预期成效的质量	发生的语境	评价语境
	社会学习作为一个框架为跨学科研究过程的评价提供了有效的模型	为评价提供模型	评价模型
	采用此标准来评价项目过程中的三个层面,个人、组织和管理者	评价项目过程的三个层面	评价对象
	例如与其他领域的研究人员交流和互动的能力	其他领域的研究人员	其他领域人员
	对个人层面学习支持的项目特征通过半结构访谈的方法进行评价	通过半结构访谈评价	访谈评价
	该框架用于探索过程、无形中介结果和有形研究结果之间的互动	框架用于探索互动	框架用途
	基于先前的研究,形成了框架中的评价过程和结果方面的标准	评价过程和结果方面的标准	评价标准
	描述了用于将评价标准进行操作化的数据集合	评价标准操作化的数据集合	操作化数据集合
de Lima 等 <sup>[49]</sup>	此评价是内部的,因为是由自项目开始就参与其中的项目协调员开展的	由全程参与的项目人员开展	参与式评价
	主要的困难在于选择、使用和解释指标	选择、使用和解释指标	指标的选择、使用和解释
	大学类似于小城市,而学生可以被认为是城市代理人	大学与大学代理人	客体代理人
	构建评价框架关注科学与技术机构的经济、环境、社会 and 机构方面的最佳表现,目的是为其他机构和社会提供示范	目的是为其他机构和社会提供示范	示范目的
Anderson 等 <sup>[46]</sup>	通过评价模型来分析科学与技术机构的社会、环境和经济影响	评价模型	评价模型
	出于此原因,使用的评价模型可以知晓学生、教师如何获取知识	评价模型	评价模型
	模型的理论立场是复杂理论	模型的理论立场	理论立场
	这种情况下形成的学习结果用于评价知识、技能、态度和行为的发展	用于评价发展	评价目的
	我们聘请了一个区域评价者,受到当地健康部门的资助	聘请区域评价者	评价者
	尤其关注所有利益相关者的观点的复杂性	利益相关者观点的复杂性	利益相关者观点
	评价方法论包括混合方法和循环式数据采集	混合方法	混合方法
	评价的结果依照 Kirkpatrick 的框架进行映射并且整理了更大范围的影响	评价结果按照框架映射	结果映射
Tackett 等 <sup>[91]</sup>	评价模型的优势在于方法的结合及不同评价工具的使用	不同评价工具	评价工具
	审计当局可以使用评价框架为医学院提供引导	评价框架为医学院提供引导	提供引导
	将评价的权值标准化可以提供基于证据的审计标准	评价权值标准化	指标标准化
Bergsmann 等 <sup>[18]</sup>	致力于定义和描述隐晦的逻辑模型来构造项目评价框架	定义和描述逻辑模型	逻辑模型
	基于评价问题说明了示例数据源和指标	基于评价问题的数据源	数据来源
	使用标准是指评价的结果必须满足潜在使用者的信息需求	评价结果满足信息需求	结果用途
	对能力的定义也不能抽象层次过低,大量的细节对潜在使用者造成负担并且与评价的灵活性标准冲突	与评价的灵活性标准冲突	灵活性标准
	与参与式评价一致,能力理论模型的评价应该容纳所有相关的利益相关者的视角	参与式评价容纳所有利益相关者	参与式评价
	利益相关者群体根据高等教育机构的不同而不同	利益相关者群体不同	利益相关者群体

续表

参考文献	原始语句	代码	概念
	有效整合不同观点的方法是形成异质焦点小组	形成异质焦点小组	访谈评价
	关于语境方面的评价考虑财政和基础设施	语境方面的评价	评价语境
	因此第三阶段的评价非常关注教学过程中的细节	评价关注细节	细节评价
Dark 等 <sup>[5]</sup>	第一步是确定评价的类型	评价类型	评价类型
	在教育干预的语境中	教育干预语境	评价语境
	形成性评价用来搜集干预过程中可以用来提升的数据和信息,总结性评价判断干预的价值	形成性、总结性评价的用途	形成性、总结性评价
	有效的评价问题可以产生有根据的、可操作的、可信的和有用的结果	评价问题产生评价结果	结果产生
	测量过程中的数据收集用于回答评价问题	数据收集用于回答评价问题	数据收集
	用不同的收集数据的方法包括调查、试验、性能测量等	不同的数据收集方法	数据收集方法
Millstone 等 <sup>[23]</sup>	参与式监测和评价要求关键利益相关者的活跃参与	关键利益相关者的参与	关键利益相关者
	影响力评价应该尽量与研究项目结合来促进内部学习过程	评价应促进内部学习过程	评价应有作用
	不同工具和方法可以搜集过程中的各个阶段的不同信息	搜集过程各阶段信息	搜集信息
Phillips-Wren 等 <sup>[89]</sup>	提出的评价框架包含组织和技术标准	框架包含组织和技术标准	组织和技术标准
	第一个理论依据是物理符号系统假说	理论依据	理论依据
	该框架是对之前评价模型的替代	框架	评价框架
	智能决策支持系统的形成和评价是为了重要的现实世界应用	评价为了现实世界应用	评价应用
Leacock 等 <sup>③</sup>	通过讨论大型评价应用来形成标准化、启发性的评价体系	标准化、启发性评价体系	评价体系
	形成的评价体系可以满足学习资源的质量保证和质量提升需求	评价满足质量保证和提升需求	满足需求
	学习目标制定和评价者都应该着重考虑隐含在学习资料中的假设	评价者	评价者
Walther 等 <sup>[40]</sup>	基于出版前评价的缺点,本文形成了6项标准	标准	评价标准
	因此它的传播还需要关注其他用户的接受程度,这决定了总体的质量	关注其他用户	其他用户
	评价由个人、期刊部、会议等共同开展	评价的共同开展者	评价开展者
	评价数据结构的标准化	评价数据结构标准化	评价数据结构
	在开放评价的语境中	开放评价语境	评价语境
Bonaccorsi 等 <sup>[81]</sup>	通过开放的科学社交网络统一框架来组织开放评价	科学社交网络统一框架	评价框架
	使用组织间理论来描述和分析大学-企业关系的两个维度	使用组织间理论	理论使用
	描述的理论框架用来理解和评价大学-企业关系的不同影响因素	用来理解和评价影响因素	评价用途
	对关系的评价需要使用定量和定性变量	定量和定性变量	评价变量
Ginsburg 等 <sup>④</sup>	对信息的搜集需要从两个机构,基于不同的机构层面和不同的时间点	从不同机构、不同时间搜集信息	搜集信息
	主治医生对医学生专业素养的评价不仅是解释性的、主观的和情境依赖性的,而且与人们对他人行为的总体判断高度一致	情境依赖	情境依赖
	主治医生对医学生专业素养的评价	主治医生	评价者
	评估的可靠性将通过关注在相互冲突的价值观背景下可观察到的行为而不是抽象的定义来提高,因为前一种方法被假定为更客观和真实	前一种方法	评价方法

注:①郭振宇. 高等教育收益评价理论研究[J]. 技术与创新管理, 2008(2): 175-176, 189;②袁安府, 陈大柔, 范柏乃. 哲学社会科学事业发展评价体系研究[J]. 科学学研究, 2008(5): 913-920;③Leacock T L, Nesbit J C. A framework for evaluating the quality of multimedia learning resources [J]. Educational Technology & Society, 2007, 10(2): 44-59;④Ginsburg S, Regehr G, Mylopoulos M, et al. From behaviours to attributions: Further concerns regarding the evaluation of professionalism[J]. Medical Education, 2009, 43(5): 414-425.

附表 2 评价实践相关成果中代码到概念的形成过程

参考文献	原始语句	代码	概念
文献 <sup>[27]</sup>	大学根据一些学术或研究绩效指标来排名	绩效指标	评价指标
	核实每个指标的数据分布是否符合事实	指标数据分布	数据分布
	在必要的时候使用标准化统计方法	标准化统计方法	统计方法
郑烨等 <sup>[55]</sup>	构建科学基金信息公开实践评价的循证评估框架	循证评估框架	评价框架



续表

参考文献	原始语句	代码	概念
	本研究中的评价主体是信息公开研究课题组	评价主体	评价主体
	最佳研究证据,强调来源于可观察到的、经过实验验证的,用于解决研究问题的各种数据和案例集	数据和案例集	数据集
	利益相关者的认知,主要包括不同的利益相关主体对科学基金信息公开实践活动的理解、态度、需求、偏好等	不同利益相关主体认知	利益相关者认知
	本研究需要厘清科学基金信息公开实践评价的具体指标及构成要素	评价的具体指标	具体指标
	基金委内部工作人员对信息公开实践的态度和认知也是构成对科学基金信息公开实践评价的重要证据来源	态度和认知是证据来源	认知评价
	本研究的理论贡献是首次尝试运用循证评估理论在科学基金信息公开实践评价领域	运用循证评估理论	理论使用
	在当前全球广泛推行信息公开的现实背景下	信息公开的现实背景	现实背景
	在发现和提出问题的同时,也为国内其他科学基金资助机构今后推动信息公开实践提供政策参考	提供政策参考	政策参考
郎永杰等 <sup>[93]</sup>	产学研合作绩效评价就是在既定的评价目标指导下	评价目标指导	评价目标
	并按照一定程序运用数理统计、运筹学等方法	数理统计、运筹学等方法	计算方法
	如产学研合作各参与者的贡献	参与者的贡献	参与者
	在实践领域,国外产学研合作绩效评价已广泛运用于政策、大学、企业和研究机构,并且已经形成规范的评估模型和评估体系	规范的评价模型	评价模型
张士强等 <sup>[56]</sup>	在评价方法上,应采用定性与定量相结合、自我评价与专家组评价相结合及模糊判断的评价方法	评价方法	评价方法
	应遵循以下原则:教学条件、教学过程与教学结果相结合的原则	遵循原则	评价原则
	通过研究,我们将每一个三级指标分为 A、B、C、D 四个等级	指标分等级	指标等级
	成立毕业设计评价专家组	评价专家组	评价专家
	分析评价结果,提出整改措施	分析评价结果	结果分析
	统计、汇总、分析各种评价资料	各种评价资料	评价资料
季靖 <sup>[85]</sup>	通过建立健全客观、合理、科学的评价体系	建立健全评价体系	评价体系
	但在实际的操作过程中,对评价结果的处理过程需慎重	评价结果的处理	结果处理
	课堂教学的实际评价没有全面而系统地进行,评价体系不够健全,也没有形成制度化、规范化	形成制度化、规范化	制度化、规范化
	组织部门不重视评价后效的追踪,不注意收集各方面的反馈信息	收集各方面反馈信息	反馈信息
	采用由学生评价、督导评价、领导评价、系部同行评价及自我评价等子系统构成的系统工程	学生评价、督导评价、领导评价、系部同行评价及自我评价	多元主体评价
	指标体系的设计必须要紧紧围绕评价目的,要体现出简明针对性	指标体系围绕评价目的	评价目的
夏云霞等 <sup>[43]</sup>	“建立合理的科研评价周期”	科研评价周期	评价周期
	实行科学合理的分类评价,按照学科领域,基础研究、应用研究等不同研究类型,建立科学合理的分类评价标准	科学合理的分类评价标准	分类标准
	在完善科研评价导向方面,要坚持服务国家需求和注重实际贡献的评价导向	科研评价导向	评价导向
	事实上,我国政府层面的评价制度尚未形成,缺乏一定的规范与适当的制度保障	缺乏规范与制度保障	规范与制度保障
	定性、定量相结合,运用各种定性及定量评价工具,量化的定量分析,同行评议与描述的定性分析	定性及定量评价工具	评价工具

续表

参考文献	原始语句	代码	概念
	各个环节都应该遵循的一个基本的原则是,评价过程是否公正、客观、公开、开放。它是影响评价质量的一个重要要素	影响评价质量	评价质量影响因素
	目标层科研团队绩效评价,从“目标实现、突出成果、人才培养、组织运行”4个维度,设立了4个“一级指标要素”,将各一级要素的内容进行分解,并结合研究所的目标导向,设立了16个“二级指标要素”,为次一级原因	科研绩效评价维度	绩效评价维度
	给每个指标要素进行赋分,通过累加得到团队得分	指标要素赋分	指标赋分
	《框架》制定了明确的价值观教育实效性评价标准	价值观教育实效性评价标准	评价标准
辛志勇等 <sup>[57]</sup>	确定更具操作性的评价指标体系	操作性评价指标体系	评价指标体系
	量性数据和质性数据同时收集	量性数据和质性数据	量性、质性数据
	他们评价价值观教育有效性采用的是一种混合式的方法途径	混合式方法途径	方法途径
王洁等 <sup>①</sup>	TALIS“教师专业发展”指标的维度构成:教师专业化的三个领域	指标的维度构成	维度构成
	通过展现大量的数据信息,帮助人们更全面地理解不同教育系统中教师的教学实践和参与专业共同体的学习情况	大量数据信息	数据信息
	TALIS教师专业发展调研框架是一个格子模型	格子模型	评价模型
	与科研评价相关的服务大致可以分为以下几个层面:①针对期刊/论文的科研评价等	科研评价服务层面	评价层面
王飒 <sup>[42]</sup>	所采用的评价指标包括:①基于论文的评价指标等	采用的评价指标	评价指标
	一些图书馆还使用SPSS、CiteSpace、HistCite、BibExcel、NVivo、NetMiner等工具	SPSS等工具	评价工具
吉汉强等 <sup>[48]</sup>	因为单项指标并不能反映评价对象(本文主要指文献资源建设)的总体水平	评价对象的总体水平	评价对象
	遵循指标体系的构建原则	指标体系构建原则	评价原则
	权重是一个相对概念,表示该指标在整体评价中的相对重要程度	权重表示指标重要程度	指标权重
	从评价结果可以看出本馆文献资源建设在哪个环节存在问题,并可以以此为依据设定一个目标值	评价结果看出问题,设定目标值	结果用途
石朝云等 <sup>②</sup>	注重多种指标进行多样性和多层次的测算,并将动态与静态指标相结合	多样性、多层次指标测算	指标测算
	建立一个稳定、规范、可操作的期刊评价体系涉及因素众多	稳定、规范、可操作的评价体系	评价体系
	本次评审的主要目的是促进系列期刊整体进步	评审的主要目的	主要目的
邱均平等 <sup>[58]</sup>	本文拟在政治历史环境、社会现实环境与学术环境三个方面进行国内外比较的基础上	政治历史环境、社会现实环境与学术环境	总体环境
	应用与发展研究成果评价的基本方法是:政策效果评估;民意测验(社会评价)和社会实验	评价的基本方法	基本方法
	这里可以对国外人文社会科学科研成果评价的数量标准、质量标准、学术影响标准、社会效益标准应用于我国的实际状况作一简要的考察分析	评价的数量、质量、学术影响、社会效益标准	评价标准
	我国还不具备有效地进行同行评议与社会公议的社会基础与制度保障	社会基础与制度保障	制度保障
胡琳等 <sup>[80]</sup>	笔者运用层次分析等数理方法	层次分析等数理方法	数理方法
	构建定性和定量相结合的图书馆员国际交流项目评价体系	定性和定量结合的评价体系	评价体系
	定量指标数据的获取	定量指标数据	指标数据
	图书馆员国际交流项目评价体系的指标覆盖了事前、事中和事后的全过程	事前、事中、事后	评价时期
文献[24]	排名公式中的指标和相应权重列于表中,相关指标列于一组	指标和相应权重	指标权重

续表

参考文献	原始语句	代码	概念
	排名中使用的数据和测量由 Clarivate Analytics InCites 提供	数据和测量的提供	数据提供
	将大学作出全球性的比较,关注大学的学术研究水平和整体声誉,为申请大学的学生提供参考	提供参考	参考作用
Granit-Dgani 等 <sup>[44]</sup>	方法论上,评价和研究需要可信、有效的工具	可信、有效的评价工具	评价工具
	本研究描述了基于理论的评价手段的形成	基于理论的评价手段	理论依据
	为教育者、评价者和研究者提供了形成性评价的有效数据	形成性评价的有效数据	形成性评价
Ladyshevsky 等 <sup>[71]</sup>	此项评价被称为 360 度评价是因为候选人受到下属、同事、上司对他们表现的反馈	下属、同事和上司的反馈	利益相关者反馈
	该数据可以直接用于课程规划、发展和评价,因为数据直接来源于工业部门	数据直接来源	数据来源
	该评价是用于发展目的而不是绩效评价	用于发展目的	发展目的
Xu 等 <sup>[94]</sup>	本文评价了网络公开学术论文的影响力,建立标准化科学评价指标	标准化科学评价指标	指标标准化
	本研究目的是构建科技论文影响评价指标,为学术资源评价提供参考	提供参考	参考作用
	学术论文的评价方法上存在争议	评价方法存在争议	评价方法
Yu 等 <sup>[50]</sup>	本研究将期刊评价指标分为三个类别	评价指标类别	评价类别
	然后计算三个一级指标的值,并给指标赋予主权重	指标赋予主权重	指标权重
	所有的数据均是标准化数据	标准化数据	标准化数据
	在该模型中,每个潜在变量具有许多相应的测量指标	模型的潜在变量	评价模型
	若只能得到指标的有限数量的数据,则无法产生正确的评价结果	数据对评价结果的影响	评价结果影响因素
Fjortoft 等 <sup>[45]</sup>	主要的评估工具是问卷调查,检查和审计	评价工具	评价工具
	使用 SPSS 进行数据分析	数据分析	数据分析
	该研究和类似评估的结果表明没有提供足够的指导来克服障碍	评价结果指导	结果用途
Wright 等 <sup>③</sup>	通过匿名调查问卷评估教育干预措施	调查问卷评价	问卷评价
	在干预前一周进行评价,然后干预后的一个月和九个月重复进行	前一周、后一个月、后九个月	评价时间
	“虽然没有计划如此,但进行的干预基本上是罗杰斯的创新扩散模型”	基于创新扩散模型	模型依据
Klautzer 等 <sup>[53]</sup>	一些评估可能发生得太早或太晚,无法捕捉研究对不断变化的政策环境的影响	不能发生太早或太晚	评价时机
	在整个研究过程中,我们应用各种数据收集方法	各种数据收集方法	数据收集方法
	回报框架适用于社会科学的研究评价	适用的评价框架	评价框架
Baschung <sup>④</sup>	对案例研究中的数据源进行了总结	总结数据源	数据来源
	收集的数据主要来源于半结构化访谈数据	数据收集方法	数据收集
	本研究选择了四个基础变量	基础变量	基础变量
Hartmeyer 等 <sup>[78]</sup>	因此数据的产生很大程度依赖访谈方法	访谈方法	访谈评价
	正规户外科学教学中的形成性评价	形成性评价	形成性评价
	探索教师对于语境决定因素的看法	语境决定因素	语境决定
Hough 等 <sup>⑤</sup>	探索产生的结果,为下一步教学的决策提供基础	结果提供决策基础	结果用途
	我们在专业发展理论的背景下进行概念化	在专业发展理论的背景下	理论背景
	采用多方面的数据收集方法来探索	多方面数据收集方法	数据收集方法
	建立可靠而有效的评价工具,为教练提供有助于他们与教师合作的信息	评价为教练提供合作信息	信息提供



续表

参考文献	原始语句	代码	概念
Duffy 等 <sup>⑥</sup>	使用 673 名心理学家的样本,计算六个生产力指数	计算生产力指数	计算指数
	那些在研究密集型环境中的人获得终身职位和晋升机会,在学术界获得更多尊重	研究密集型环境	研究环境
	研究人员主要使用生产率测量来评价特定领域内人员,计划和机构的研究生产率,以及与个人生产相关的变量	生产率测量变量	测量变量

注:①王洁,张民选. TALIS 教师专业发展评价框架的实践与思考——基于 TALIS2013 上海调查结果分析[J]. 全球教育展望, 2016, 45(6): 86-98;②石朝云,游苏宁,蔡丽枫,等. 医学期刊综合质量评价的实践与思考[J]. 中国科技期刊研究, 2008, 19(6): 941-944;③Wright I M, Wake C H, Anderson H, et al. Assessment of the multidisciplinary education for a major change in clinical practice: A prospective cohort study[J]. BMC Health Services Research, 2009, 9(1): 28;④Baschung L. Identifying, characterising and assessing new practices in doctoral education[J]. European Journal of Education, 2016, 51(4): 522-534;⑤Hough H J, Kerbow D, Bryk A S, et al. Assessing teacher practice and development: The case of comprehensive literacy instruction[J]. School Effectiveness and School Improvement, 2013, 24(4): 452-485;⑥Duffy R D, Jadidian A, Webster G D, et al. The research productivity of academic psychologists: Assessment, trends, and best practice recommendations[J]. Scientometrics, 2011, 89(1): 207-227.