

# 中文人文社会科学学术图书 质量评价体系再探讨

◎ 叶继元

**摘要:** 针对目前我国大规模人文社科学术图书质量评价缺失、同行评价不足、评价指标体系与评价目的不匹配等问题,运用学术“全评价”体系分析框架,通过比较研究法,在吸取已有研究成果的基础上,从评价主体、评价方法、评价制度等方面,构建了以学术图书质量评价指标体系框架为核心的“全评价”体系及其方案。

**关键词:** 学术图书;人文社会科学;质量评价;评价体系

**课题:** 国家社会科学基金重大项目“新时代我国文献信息资源保障体系重构研究”(19ZDA346)

**DOI:** 10.3969/j.issn.2095-0330.2020.05.006

学术图书与学术论文是研究成果最主要的两种类型。一般说来,论文反映的是最前沿、最尖端的研究方向和成果,新颖性和前沿性是学术论文的主要特色,而内容的相对系统性、完整性和理论性则是学术图书的主要特点。相对于自然科学领域更重视期刊论文而言,社会科学尤其是人文学科领域更重视学术图书,因为学术图书是对一个专题的较为系统的阐述,通过对图书尤其是著作或专著的阅读,有助于比较深入和全面地了解某一专题的研究成果,对其前因后果、来龙去脉、逻辑关系等有更细致的掌握,有利于以此为基础进行学术创新,促进知识不断增长。因此公平、公正、公开和科学地评价学术图书的质量,是当前学术评价中一项极为重要的任务。

## 一、评价背景与意义

什么是学术图书、学术著作和学术专著,笔者曾有过专文论述,<sup>①</sup>在此不赘,仅引述要点。

学术图书(Academic/Scholarly Book)是指内容涉及某学科或某专业领域,具有一定创新性,对专业学习、研究具有价值的图书,通常在书中有文献注释或参考文献,书后有索引。它包括学术著作、学术专著、

学术论文汇编/论文集、会议录、大学及以上程度的教材/教科书和参考工具书、学术随笔等。其“著作方式”多数是“著”“撰”,少数是“编著”(编著是一种著作方式,汇集其他多个作者、多种作品的思想、观点和内容资料,但有作者自己独特的见解陈述,凡无独特见解陈述的书稿,不应判定为编著),工具书、大学教科书则是“编”。但是中专及以下层次的教材、通俗读物、时事读物,以及一般的字典、词典等不应包括在内。

学术著作(Academic/Scholarly Writing, Works)是指以问题或专题为中心,具有创新性和逻辑性,能自圆其说的学术图书。它包括学术专著、学术进展评论、著作性研究指南、手册等。其“著作方式”多数是“著”“撰”,少数是“编著”,一般没有“编”。

学术专著(Academic/Scholarly Monograph)是对某一学科、领域、专题进行较为集中、系统、全面、深入论述的著作。一般是对特定问题有独到见解,且大多是“自成体系”的单著或二三人合著的学术著作。它包括单本专著、多卷集专著、专著丛书等。与学术论文相比,学术专著的篇幅较大,内容所涉及的问题一般也较专深,更具专业性、系统性、全面性、深入性,论述或论证具有广度和深度。撰写人一般是单独作者或几个作者,一般不会像学术论文那样一份成果作品的作者署

名多到数十个或数百、数千个。与学术著作相比，学术专著的创新层次较高，一般具有原创性，而学术著作中有许多著作属集成创新或应用创新。学术专著的“著作方式”都是“著”“撰”，没有“编”或“编著”。

然而有一种现象值得注意，就是有一些图书的形式和内容不一致。从内容上看是学术著作或专著，从形式上看则是教科书（为了容易出版或扩大读者面）；而有些书则相反，实际上是教材，却以学术著作或专著的形式出现。一般来说，学术性教材是学术图书，但不是学术著作，更不是学术专著。同理，学术性随笔、学术性工具书、学术文章汇编和精选等是学术图书，但不是学术著作或专著。由此可知，虽然对学术著作的内涵和外延有较为明确的界定，对大多数出版物能够很快确定其是何性质，但对混合型出版物应慎重认定，根据评价目的对学术图书适当分类。对这种“亦此（学术著作）亦彼（学术图书）”或“非此非彼”的图书，在进行“分类”时应格外留心，应重“实”不唯“形”，以免张冠李戴。

学术图书是学术活动和研究成果的重要呈现载体，是社会发展尤其是创新型国家所需要的重要信息资源。人文社会科学和自然科学是“鸟之两翼”，缺一不可。从人文社会科学撰写和发表特点看，人文社会科学特别是人文学者更重视图书型的成果。近20年来，中国经济有了很大发展，相应地，社会、文化和学术亦有大发展。随着研究者数量、科研经费数量、出版社接受书稿数量的增加，学术图书的品种类型、数量亦有很大增加。几年前，笔者曾对学术图书、学术著作和专著的数量做过考证和推测。据保守统计，2013年我国出版图书444 427种，其中初版（新书）255 981种，初版学术图书76 794种。在初版学术图书中，学术著作38 397种，学术专著25 598种。其中文科学术图书44 464种，文科学术著作22 232种，文科学术专著14 821种。<sup>②</sup> 2018年我国出版图书519 250种，其中初版（新书）247 108种。<sup>③</sup>与2013年相比，2018年出版图书总品种数多了7万多，但新书却少了0.8万种。据此推算，2018年至2020年出版的人文社科新版学术图书、著作和专著与2013年大体相当，恐略有下降，因为目前学术图书的出版费用大涨。总体看，中国已成为学术图书出版大国。学术图书数量上去了，就质量来说有些

书提高了，但整体质量似乎并没有得到相应提高，学界、出版界批评之声甚多，2020年中有关《平安经》一书的出版及对“明学”的虚假评价即是典型案例。为了能确切了解中文人文社科学术图书的质量和水平，拿出具有说服力的数据、事实，有必要对学术图书进行较为全面、系统、科学权威和公开透明的评价。这对于提高学术图书出版质量、倡导全民阅读、提高民众的文明程度和我国软实力、促进优秀图书成果“走出去”，对于图书作者、学术机构/大学、出版社的评价，对于学术资源的合理、优化配置等，都具有十分重要的意义。

## 二、国内外相关学术图书评价综述

十多年来，出版界、图书情报界、学术界、教育界、媒体和科研管理部门均进行过图书评价，较有影响的有以下一些项目。

一是管理部门的评价，诸如在中宣部、新闻出版总署指导下，由中国出版工作者协会主办的中国图书奖评价；教育部人文社科优秀成果评价中包括图书评奖；国家社科规划办国家哲学社会科学成果文库评价；各省市规划办社科成果评价。二是大学、研究机构的评价，诸如中国人民大学吴玉章奖评价，社科院孙冶方经济科学奖著作奖、胡绳史学图书奖评价，华东师大思勉人文学科原创作品奖评价。三是图书情报机构的评价，诸如国家图书馆文津图书奖评价，中国图书馆学会信息资源专业委员会、北京大学图书馆有关“核心图书”评价，2005—2007年由教育部高等学校图书情报工作指导委员会、中国图书馆学会高等学校图书馆分会立项的课题——“中文图书评价研究”，上海师大图书馆《中国哲学社会科学类学术图书基本书目（1995—2005）》评价，南京大学、中国图书评论学会“中文学术图书引文索引”数据库评价。四是学会、媒体等的评价，诸如中国图书评论学会“中国好书”评价，《光明日报》“光明书榜”评价，《中华读书报》“月度好书榜”评价，百道网“中国好书榜”评价，商务印书馆2013年开始评选“年度十大好书”（包括中译本，由媒体人、书评人、学科专家及各高校阅读团体网上评选，以后30多家出版社跟进评选），北京人天书店有限

公司《中国核心书目》的编制等。

以上这些评价,评价目的不同,评价主体、评价标准及指标也不同,评价的方法、制度设计亦有差别,其影响力各异。总体看,图书情报界侧重于文献计量方法。北京大学图书馆有关“核心图书”的评价,对1949年以来大陆出版的100多万种中文图书(含译著)进行分学科评价,从经典性、学术价值、使用价值等多种角度着手,遴选出5%左右学科学术影响力大、学术价值高的核心图书。该项目是迄今为止最为认真和系统的一次近现代中文图书整理工作,但由于缺乏相应的工具性数据库,该项目主要通过人工来完成图书的被引量、借阅量、被摘量的统计,在数据采集及整理方面工作量很大,持续了相当长时间。此次评价主要是为收藏图书,与图书质量评价有一定的联系,但并不是图书质量评价。2015年南京大学中国社会科学评价中心筛选“中文学术图书引文索引”来源图书(引文数据统计源),则采取定性与定量相结合的方法,精选出数千本来源学术图书,其目的是从图书引用的角度,考察其被图书、期刊论文等引用的情况。显然,从对图书质量评价方面来看,这比北京大学图书馆“核心图书”的筛选更进了一步,引文数据反映的是同行专家/作者之行为,与同行专家专门对图书质量进行评价具有一定的相关性,但仍不能完全等同视之。而管理部门、学术机构的评价主要采用同行专家评价的方式,如思勉人文学科原创作品奖评价,其从图书推荐到专家评审再到评价答辩,均采用同行专家/学者评价,影响力较大,但评价的图书数量很少。尽管这些评价的图书范围、品种等有限,但无疑对于在较大规模上开展学术图书质量评价具有重要参考价值。

国外主要通过书评和评奖、畅销书榜等对图书进行评价。书评由媒体选择有声誉、有水平的书评家撰写,一般具有较高的可信度。评奖则由同行专家学者主导,影响力较大。当前可以相对便利地获取美国ISI(the Information Sciences Institute, 现属科睿唯安集团)“图书引文索引,社会科学卷”(BKCI, SS)和南京大学“中文学术图书引文索引”(CBKCI)引文数据,通过结合以图书为统计源的引文数据与以期刊为统计源的引文数据,对图书展开综合评价成为可能。

### 三、“全评价”分析框架应用于学术图书质量评价的可能性分析

“全评价”体系分析框架是笔者与团队于2010年初正式提出的有关学术评价理论的成果。

首先,“全评价”体系分析框架是根据学术评价的共性,尤其是人文社科评价的个性来构建的。它从丰富的评价实践出发,归纳概括出该体系至少由评价主体、客体、目的、方法、标准及指标、制度六大要素组成。在评价客体中,图书与期刊、学者、机构等一样,是其重要的组成部分。因此,在适用于整个人文社科评价客体的通用框架下,结合学术图书的特点,就能构建成学术图书质量评价体系。

其次,质量评价主要侧重于对评价客体内容、效用的评价,对评价客体形式、数量的评价则相对淡化。而“全评价”体系分析框架提出的三大维度——形式评价、内容评价、效用评价正好符合质量评价的要求,即注重内容与效用评价。

再次,“全评价”体系分析框架非常强调评价目的的制约作用。“质量”评价的目的不同于一般的量化评价,而“质量”本身也可以有不同的目的。目的不同,在评价主体、客体、方法、指标、制度等设计上都不一样。根据不同的评价目的,可以有多元评价标准。评价客体的分类亦如此。之后会举例说明。

最后,“全评价”体系分析框架也非常强调学术共同体,尤其是小同行专家评价的作用,强调对同行评议专家局限性的消除要注重公开、申诉、元评价等评价规章制度建设,同时不忽视引文等计量数据的应有作用,摆正其与同行专家评价的关系。比如质量评价的目的是优中评优的评价,还是基本合格的评价,不同的评价目的决定选择什么样的评价主体,如果是全国性优中评优、有名额限制的评价,就要选择国内外一流同行专家学者作为评价主体,评价结果的公示、申诉、复议,以及评审专家的信度考察等程序与制度就应着重设计和实施。评价主体中的同行专家学者是评价的主导,评价制度则是评价的基础和保障。目前国内许多评价之所以没有达到奖勤罚懒、褒优贬劣、激浊扬清、调动研究人员和评价专家积极性的目的,其中一个主要原因,就是没有明确评价目的,并根据评价目的来选择评价主



体、方法等。近期中央有关文件倡导代表作评价、同行评议,反对“唯SCI”,引发学界热议。“唯SCI”的要害是绝对化、片面化,理当摒弃;不“唯SCI”而合理利用SCI,则理所应当。同行评议固有的局限性,也应用制度来制约,用引文等数据来验证或补充。

2013年,笔者指导博士生王兰敬将笔者提出的“全评价”体系应用于人文社会科学学术图书评价领域,拓展了学术图书质量评价框架体系,建立了中文人文社会科学学术图书“全评价”模型。根据形式评价、内容评价和效用评价的新概念组合,采用引文等文献计量评价与同行专家评价相结合、以同行专家评价为主导的评价方法,基于人文社会科学学术图书“全评价”模型筛选出高质量的学术图书“代表作”,再通过学科同行专家小组审核方能确定为精品图书,即作为统计源(来源图书)构建中文人文社会科学学术图书引文索引示范数据库。<sup>④</sup>这说明该框架应用于学术图书评价具有可能性。

## 四、人文社科学术图书质量评价体系

根据“全评价”分析框架,结合我国学术图书的出版情况,笔者对图书出版后的质量评价体系及其方案有如下思考。

### 1. 评价目的

评价可以促进学术图书质量的提高和学术的健康发展。详细知晓近年来人文社科各学科学术图书,尤其是原创新学术著作或专著的综合质量状况,其中优秀、良好、合格(达到正式出版水平,有价值)和不合格(未达到正式出版水平)学术图书的数量、比例、作者、出版社、地区分布等情况,可以为提高学术图书的选题、编辑加工、出版、选购、阅读、利用等的水平提供有价值的参考,为合理、全面评价各学术机构、大学的质量提供参考。

### 2. 评价主体

以同行专家,尤其是小同行专家评价为主,通过专家审读文本、参阅有关引文数据、评价数据和其他评价信息,独立得出评价结论。同行专家包括该学科

或领域的研究者、学者、教授、编辑、出版者、核心读者、媒体学科专家、科研主管、硕博士研究生等。

大规模学术图书质量的评价完全采用同行评议方法是非常困难、复杂的,从某种角度上看,仅仅依靠某个机构来承担,可能有些力不从心、力所不逮,还需要有国家层面的支持和较为长远的政策保障,以及基础工作的长期积累和准备。图书质量评价的组织者非常重要,建议由学术共同体的代表,即学术水平高、公信力高、对学术评价有丰富经验的人文社科学者组成评价班子,对评价指标的设计、方法的采用、同行专家的选择,以及评价程序、评审意见、评价结果的公开等各项事宜有决定性的权责,科研管理机构或有重要影响力的学会、协会、民间团体可以提供指导或物力支持和服务,几个重要的评价机构可以提供有关数据或承办具体事宜,彻底改变学术圈外人员或单位组织学术评价的弊端,回归专业的事由专业的人来办的原则。

### 3. 评价客体

前文已提到学术图书的具体所指。第一阶段可对2015年出版的一万多种学术专著或某个学科的千本专著进行示范评价,取得经验后再逐步推开。为了提高操作性,可以采取出版社推荐、作者自荐方式,对拟参加评价的专著设置一些必要的门槛,如必须为有声誉的出版社出版的学术图书、高被引图书、有高质量书评的图书、著名学者的学术著作、被翻译成外文的图书、被转载摘编较多的图书、读者利用较多的图书、各著名高校教授为研究列出的经典参考图书、著名专家推荐的图书等,以提高评价效率。取得经验后,再对学术著作、学术图书进行质量评价。

人文社会科学是以人、人化的自然和社会为研究对象的,其研究成果无法通过实验重复验证,其学术价值常常并不是即时性显现的,有时甚至是相当滞后的。一些极具价值的理论、思想或观点,是在经过了长期社会实践之后,才为人们所认识和发掘的,有时甚至可能产生与当时认知相反的颠覆性的价值判断。

对于人文社会科学的研究成果,一个观点正确与否,需要时间的检验,才能完全认识到它的意义和价值。历史评价和社会评价或曰实践才是学术研究成果的终极评价,从某种意义上说只能通过时间的检验才

能得以证明。时间或称历史判断对于人文社会科学研究成果尤为重要,因此被评价的图书一般要求已出版5年以上。

遵循“分类评价”的原则。人文社会科学包含不同的学科、领域,学术图书的类型多种多样,需要多元化评价和分类评价。只有按类评价,才具有横向可比性,评价结果才更有意义。

#### 4. 评价方法

对各出版社出版的学术图书,尤其是学术著作或专著的质量进行“出版后”的评价。以同行评议法(peer review)为主,辅以文献计量法。文献计量数据,诸如引用次数、转载量、文摘率、被书评或评论情况、获奖情况等仅提供给评审专家参考。文献计量数据单独加权计算,不与专家评议结果综合加权,评议专家意见独立加权计算。图书引文索引数据库可加以利用。

要在评价的不同阶段采用不同的评价方法,以充分发挥定量方法和同行定性评价方法的作用,将定量与定性方法有机结合起来,但不搞简单化的“平衡”。在充分考虑人文社会科学特征的基础上,也要合理应用目前成熟的定量评价方法。仅仅通过专家推荐点评,工作量极大,且具有某些随意性和不稳定性。定性评价与专家的专业水准、学术能力、个人志趣和喜好等有着极大的关系,而这些因素都属于专家的隐性知识,是不确定的,也是难以把握的。因此,要注重价值性判定与客观性判定相结合。在初级评估阶段,以定量评价为主,专家可对例外情况进行补充;在评价高级阶段,则以专家评价为主,即尽可能多地听取本研究领域,甚至外领域众多专家、学者的评价意见,并可用引文等定量评价校对、验证。

#### 5. 评价制度

建立同行专家动态遴选制度、考核制度、淘汰制度、增补制度,评价流程监督制度,评价结果公示制度,评价对象申诉制度、反馈制度,评价结果共享制度等。质量评价主要依赖于同行专家评价,对专家评价的局限和弊端需要通过评价制度的建立和完善来尽量减少。

评价专家库由相当数量的国内外专家组成,这些专家可由相关学科学会、协会,或著名学者推荐产生,也可利用全文、引文等各种数据库对研究成果突出、影响力大的作者进行筛选,力争将近年来仍然活跃的真正的同行专家组织起来。评价专家考核制度主要考核专家的认真积极程度和诚信情况,充分尊重专家评审意见的知识产权和工作业绩,对认真的专家一定要给予必要的精神与物质奖励。对于那些诚信有问题、不认真参加考核或在考核过程中出现故意偏袒、刻意针对某些图书等有不良行为的专家执行严格的淘汰制度,并规定在若干年内不再聘任。专家库是动态的,在严格淘汰制度的同时,需要及时补充新鲜血液,建议专家库每年进行动态调整。评价对象和评价结果通过各种媒体及时向学界公示,征询更广范围内学者的反馈意见,同时接受被评价对象的反馈和申诉,以确保评价结果公正合理。专家的评价结果可整理成评价报告,提供给出版社作为参考,供其在出版过程中借鉴,以发扬优势、规避短处,从而有利于学术出版的长远发展。

### 五、评价标准及指标体系框架

为了更好地阐述学术图书的评价标准及指标体系框架,下面先讨论人文社会科学成果评价的标准及指标,再提出学术图书评价标准及指标体系框架。

所谓评价标准,是指人们在评价活动中应用于对象的价值尺度和界限。评价的客观性因素是评价标准具有科学性的重要依据。评价标准是评价活动的关键、核心部分,是人们价值认识的反映,它表明人们重视什么、忽视什么,具有引导被评价者的作用。评价的标准依赖于评价目的,即评价目的决定评价标准。评价理论性成果与评价应用性成果的标准就不一样。

“在学术评价中,科学家用真和美的(标准)尺度去衡量不同的科学理论的价值,对它们做出比较和选择。哥白尼理论比托勒密体系具有更大的数学简单性和和谐感。评价科学理论的标准应该是:(1)理论的可证伪程度较高;(2)理论经受检验的严峻程度较高。”<sup>⑤</sup>应用性成果则以应用理论、社会效益、社会舆论和风气为标准。社会科学领域,尤其是人文学科领域中的

价值判断，一定程度上是见仁见智的事情，很难建立起一个公认的客观标准。但价值判断必须以事实认定为前提，只要证据充分，可以做到相对客观、公正，就可以认可其属于“事实认定”。以事实为根据，以学术共同体应有的价值观为指导，不同的同行专家应该可以获得相对的评价共识。

社会科学与人文学科有其不同于自然科学的特性，故其评价标准也有别于自然科学的评价标准，即以社会效用（价值）作为其最高标准。从这个意义上看，形式评价、内容评价和效用评价既可以是评价方法，也可以是评价标准。形式评价与内容评价实际上可以被看成是一种论证性评价。此类评价的标准是逻辑标准，即评价对象是否得到学术界的认同及其所获认同程度的高低。效用（价值）评价也可以被看成经验性评价。此类评价标准是事实标准，具体表现为评价对象的实践情况。而人文学科的实践主要是指社会舆论，尤其是人文学科学术共同体的舆论；社会科学的实践主要是组织实践，主要是决策部门主导的实践。<sup>⑥</sup>

如果说评价标准是总的衡量尺度，那么评价指标可以看成是评价标准的细化。每一评价指标都是对评价对象某个或某些属性的概括，作为评价指标，它必须在质和量两方面都可以确定。比如，引文量、获奖数等，既有明确的含义，又可以量化。有些指标有明确的含义，但短时间内很难量化，如声誉、创新程度、对学科的贡献度等，只能通过调查事件、案例或专家根据事实、数据等直接用程度词来描述。这也就是内容评价和效用评价费时费力的一大原因。对于定性的指标，如评价研究成果，一般可从准确/科学性/规范性、价值性、创新性三大方面设置指标。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；评价指标体系在逻辑上是自洽的，不宜太多或太少。首先要增加指标体系的民主性和公开性，应经广大同行民主程序的广泛讨论，征求意见。其次要增加指标体系的科学性，要由专家充分论证指标是否能反映评价目的，是否有遗漏的指标，以及各个指标在体系中所占的权重是否得当。最后，指标体系要做到相对统一性与差异性相结合。统一性是指在条件相同的情况下，指标体系应该可以做适当的比较；差异性是指考虑文科各学科或领域的差异，针对具体

情况，对指标做出调整。要在有统一指标框架的基础上体现不同类型、不同学科或领域的不同指标。

我们认为，人文社会科学评价指标的选择应该遵循如下原则：一是指标应与评价目标有关，随评价目标的不同而不同。不同的评价目标，应该选用不同的评价指标，因此人文社会科学评价指标应该是一个动态的评价指标体系，但它们遵守一个共同的指标原则，随评价目标的变化而变化。二是指标组合可以形成一个多方位的立体评价维度，避免盲人摸象的嫌疑。三是与评价指标相关的数据应易于收集。如果设置了相关指标，但无法获取相关数据，所选取的指标理论上可以保留，但实践中则暂不采用。

笔者及其团队曾对全国性的人文社会科学专家展开有关成果评价看法的调查，调查对象是教育部社会科学委员会委员，共发出调查问卷110份，回收问卷109份，有效问卷101份。统计结果表明，45%的专家支持对成果采用专家定性评价方法；35%的专家支持通过被同行引用情况进行评价；27.5%的专家认为目前的人文社会科学评价应该定性与定量相结合；9%的专家支持以定量为主的评价；58%的专家认为人文社会科学评价应该定性定量相结合，但要根据不同情况以定性或定量为主评价，并根据各指标的权重来测量。

调查结果表明，对于人文社会科学成果评价，一般的形式评价指标体系是有缺陷的，应该建立内容评价指标体系，这是一个以同行专家为主的评价指标体系，适当兼顾定量指标。

为了模拟建立人文社会科学研究成果的内容评价体系，本文选取社会科学研究成果中的基础研究成果为评价对象，并对社会科学研究成果评价指标中的可能指标进行了初步汇总和归纳（见表1）。

表1 社会科学研究学术图书评价指标汇总

编号	可能的指标名称及含义
1	选题难度(成果中涉及的研究对象的复杂性或者难度)
	选题的前沿性(研究选题符合学科发展趋势)
	选题的价值性或现实意义
2	资料的客观性(成果中所依托的资料或数据符合客观事实)
3	学术引证规范性(成果引文符合学术引证规范)



续表

编号	可能的指标名称及含义
4	成果的出版层次（刊物的级别或出版社的级别）
5	成果的获奖级别
6	论证的逻辑性(研究过程中概念明确、严谨，逻辑论证严密，方法可靠并且有效)
7	成果科学性(成果结论明确，符合客观事物的发展规律)
8	成果创造性(成果中的理论、观点、方法和建议方案有创新性，对学科发展和建设有贡献)
9	成果的时代性(成果具有鲜明的时代特色)
10	成果实用性(成果对社会实践具有引导作用，提出的方案、理论、方针可行)
11	成果的价值（成果具有社会价值和潜在的经济价值）
12	成果的总被引用次数（含自引）
13	成果的被引系数(总被引次数与成果所属学科平均被引次数之比)
14	成果的他引次数
15	被利用次数（复印率、点击率、下载率、转载率）
16	被有评价作用的二、三次文献（检索期刊、评论期刊）收录和被评论次数
17	成果的他引系数(他引次数与成果所属学科平均他引次数之比)
18	成果学科影响指数(成果被其他刊物引用的期刊数量)
19	成果学科扩散指数(成果被本学科外其他学科引用总次数与该学科成果平均被引用次数之比)
20	成果即年影响指标(成果发表当年被引用的次数与该学科成果当年发表总数之比)
21	成果的半衰期(成果的生命力指数)

根据笔者所建立的指标框架，以上指标可以被归属为选题（方向）重要性、方法（过程）科学性、成果（结论）独创性三个方面。其中与选题（方向）重要性相关的指标有：选题的复杂性与难度、选题的前沿性、选题的现实意义。与方法（过程）科学性相关的指标有：资料的客观性、论证的逻辑性、学术引用的规范程度、成果的科学性等。上述列表中的其他指标属于从不同视角对成果（结论）的独创性进行验证的指标，主要从成果的内在价值和外显价值进行判别。成果的内在价值必须由同行专家评议做出判断。外显价值属于一种社会价值，可以看成是其产生的影响，而这影响还有两层含义：横向影响，指当下的影响面；纵向影响，指成果影响的历史跨度即半衰期。人文社会科学研究成果产生的影响在社会价值方面表现尤为明显，同时不少成果在一定程度上也能促进经济价值的实现。

基于以上分析，对上述指标进行整理，对一些具

有近似内涵的指标进行合并，模拟构建了既包含定量指标又包含定性指标的社会科学研究成果内容评价指标体系（见表2）。

表2 研究成果内容评价指标

一级指标	二级指标	指标说明
选题重要性	前沿性	选题符合学科发展趋势或具有鲜明时代特色
	价值性	选题对社会实践具有引导意义
	复杂性	成果中涉及的研究对象的复杂性或者难度
	系统性	成果中涉及的研究对象的完备程度
方法科学性	资料或数据的客观性（资料质量）	成果中所依托的资料或数据采集方法可靠，来源客观
	学术写作与引证规范性（规范质量）	研究内容重点突出，思路清晰，引文符合学术引证规范
	论证的自洽性（研究过程质量）	研究过程中概念明确、严谨，逻辑论证严密，能自圆其说
成果独创性与影响力	成果的创新性	成果中的理论、观点、方法和建议方案有创新性，创新程度的大小，对学科发展和建设有贡献
	成果的被引系数	总被引次数与成果所属学科平均被引次数之比

为了便于记忆和传播，可以将学术成果等的评价指标概括成“三面三力（度）”，亦可称“三三制”，即形式、内容、效用三面，大体对应资力、实力、效力三力（见表3）。每个“力”下一般可再分为若干子“力”。比如资力可分成支撑力、学术素养力和生产能力及影响力等；实力可分成创新力、准确力和解释力等；效力则是思想启发力、物质贡献力和精神文明推动力等。三面是所有评价客体所具有的共性，而“力”则允许有个性、有变通，且指标的解释也可不同。影响力是指以前的研究积累、知名度、被引、被摘、被评论、自我提升能力等情况。以上这些指标都是形式指标，只说明具有“资力”，是指具备了一些基本条件，拥有一些资源，在一定程度上可以反映内容的大概情况，但远不能代替内容。

内容—实力—创新力、准确力、解释力等是反映实际能力的一组指标。创新是“有中生新”，没有知识积累、根基，没有科学性的所谓创新不是创新。学术研究的本质是求真，是发现事物的本质和规律或真谛，而大大小小的创新是探索本质和规律道路上的体现，因此是可以测度或评价的。创新力可以指原始创新、集成创新、引进吸收消化再创新和移植创新能力，也可以分成选题创新、方法创新、观点创新、理论创新、资料创新、论证创新、结果结论创新的程度。准

确力可以指概念、逻辑、理论、方法、资料引用等的合理、规范程度,也可以反映研究的难度和复杂度。解释力主要指表达能力、文字水平、清晰程度。对于图书而言,创新力主要体现在选题、方法、观点、资料数据、论证、结果与结论新颖性上面。而当评价客体为研究人员时,则是指观察力(阅读力、发现问题能力、思考力)、分析力等。

效用—效力—精神启示力、物质贡献力、文明推动力对所有评价客体均有很高的概括性。精神启示力是指对人们思想上的启发、影响,或曰“社会效益”,可通过解读各种阅读数据、被引数据和事实等反映出来。物质贡献力是指经济效益,常用金额反映。文明推动力则包括精神和物质两个方面,常指二者融合在一起的情况。

表3 人文社会科学研究成果评价指标体系(FCA)

面	一级指标	二级指标	指标含义
形式	资力(指经济力量、财力等)	支撑力	经费、办公、网络、人员等基本条件
		学术素养力	学位、学历、职称、人员结构、学科优势等情况
		生产及影响力	研究积累、知名度、被引、被摘、被评论、自我提升能力等情况
内容	实力	创新力	原始、集成、吸收消化、移植创新,或选题、方法、观点、资料数据、论证、结果结论等创新
		准确力	概念、逻辑、理论、方法、资料引用等的合理、规范程度,也可以反映研究的难度和复杂度
		解释力	说理与预见程度、表达能力、文字水平、清晰程度
效用	效力	思想启示力	对人们思想上的启发、影响
		物质贡献力	经济效益,常用金额反映
		文明推动力	社会效益,对文化的贡献

实际上,许多指标是非线性的,例如引用次数、获奖次数,既是形式评价指标,也可以作为效用评价指标,主要供同行专家参考、分析之用。但为了便于操作和记忆,归纳成“三面三力(度)”指标体系。一般说来,目前国内外各种评价客体的众多指标均可以纳入进来,当然,在实际应用时切忌生搬硬套,允许根据实际评价对象有些许变通。

从评价标准及指标对其他要素的影响看,评价标准及指标是联结评价主体与评价客体的纽带,是评价得以进行的核心要素。同时指标对评价主体的学术水

平有不同要求,比如,如是出版图书数、引文数、教授人数等形式评价的一些指标,则对评价主体的要求不高,所花时间亦不多;如是内容或效用的一些指标,则对评价主体的要求很严,所花时间很多。同样指标对评价目的也有反作用,比如有的指标可以反映评价目的,但目前有关这些指标的信息收集不到或收集很不全,则要修改评价目的,目的一修改,则相应的评价方法、评价制度等也要修改。

从评价标准及指标的共性与个性看,任何事物都具有共性和个性,评价标准及指标亦不例外。所谓共性指标是指某大类评价客体都具有的特点,比如创新性,这应该是所有学术论著的共性指标,不论论文、著作还是研究报告,因为没有创新就不是“学术”论著。但具体是资料创新,还是研究方法创新,这相对来说就是个性指标,因为对于史学学术论著来说,资料创新是常态,但对于哲学而言则不是。以此类推,个性指标是某一小类评价客体所具有的特点。因此,在实际评价中善于区分共性指标与个性指标,不以共性指标代替个性指标,也不以个性指标排斥共性指标,具有重要的意义的。

本方案根据“全评价”分析框架,结合学术图书的特点,提出以形式、内容、效用相结合的评价指标体系框架,将文献计量法与专家评价法结合起来,以专家评价法为主,力求简明清晰,使之具有代表性和较高的可操作性。学术图书质量评价指标体系框架见表4。

表4 学术图书质量评价指标体系框架及其内涵

一级指标	二级指标	三级指标及评价要点
图书形式	资源形式	版本:符合出版标准规范、校对质量、印装质量,著作方式一般为“著”“撰”等;装帧设计;核心作者(作者是否为该学科专业领域的核心作者);核心研究机构(作者所在研究机构是否为该学科领域的高水平机构)。由数据和专家评价。
	影响形式	作者、作者所属机构、出版社声誉、品牌度(是否为高水平的核心出版社);是否再版、重印次数等;图书被借阅量、被国内外著名图书馆收藏数、发行数、图书被引次数、书评篇数、书评文摘量、被评论次数、获奖量/层次、被译外语种数、被摘要次数、被转载量等情况。由数据和专家评价。
出版内容	创新力(程度)	原始创新、集成创新、吸收消化移植创新,或选题、方法、观点、资料数据、论证、结果结论等创新。由专家评价。
	准确力(程度)	概念、逻辑、理论、方法、资料引用等的合理、科学、规范程度,也可以反映研究的难度和复杂度。由专家评价。



续表

一级指标	二级指标	三级指标及评价要点
出版内容	解释力 (程度)	表达能力、文字水平、清晰程度。 由专家评价。
出版效用	理论效用	对学科发展、理论认识、思维与研究方式的作用和影响；对社会、民生、道德文化建设的舆论导向作用和影响(可用案例、事件及引文次数、文摘次数、altmetrics, 下载次数、被评论次数、获奖次数等说明)。由专家评价。
	实践效用	对人才培养、社会实践和政治、经济发展具有导引、推动作用或潜在作用(主管部门的批复、采用等证明材料情况)。由专家评价。

上述指标对于学术图书质量评价的重要性以及可获得性是不一样的,其重要性不尽相同,权重的确定可采用德尔菲法,通过对学科专家、图书编辑、科研主管、研究生群体等调查后分析获得。

在上述指标体系中,并非所有指标都是可获得的,例如对于学术图书的引用指标可获取一部分,书评指标的获取就十分分散、不完整,对于学术图书的理论效用也是难以准确评估的;相对容易获得的借阅量指标,“噪音”不少,干扰因素非常多,受到借阅者自身学术水平、曲高和寡等多种人为因素的影响。所以从理想状况来看,以上指标都为必备,但事实上所有指标都完备的情况是很少的,多数情况是指标有缺损的,因此本研究认为在评价某一本、某一类或者某一时段的学术图书时,可以根据具体的评价目的进行有针对性的指标选择,并给予相应的重要性权重分配,从而建立一个相对合理的量身定制的图书质量评价方案。

值得强调指出的是,对于一本学术图书来说,作者、出版社和研究机构的知名度或品牌这三个评价指标覆盖了形式指标,某种程度上反映了一定的内容,同时又能够确实一定程度上反映出一本书的学术质量。这三个指标是容易获得的。优秀的作者不一定每本著作都优秀,但优质的著作必然是优秀的作者撰写的;同样,研究机构表明是某个研究方向的一个研究者置身于一个高水平的研究机构,从某种意义上已经体现了同行之间的学术评议;同理,某个学科领域高水平的出版社,对于图书的出版,必定也已经从选题的筛选、编辑、校勘等多个环节进行了质量控制。当然,这些评价属于形式评价,如果是重大、精细的评价,还是要同行专家对图书内容及效用进行审读性的精细评价。精细评价的程度可以根据评价目的及人力

物力投入情况而定。为弥补形式评价评出的候选图书名单之不足,多位专家联名推荐的学术图书亦可作为补充参考。

以上根据“全评价”分析框架构建了人文社科学术图书质量评价体系及其实施方案,这是一个通用的体系框架,如果要实施,还要对六大要素的各个细节进行深化,也可借助网络、人工智能、区块链等新技术,建立同行专家遴选、评审意见公开、精英同行与大众同行互动、专家评价绩效登记等一系列评价制度和评价机制,逐步完善学术图书质量评价体系。

六、结语

学术图书是人文社会科学研究的重要成果类型,对其进行科学权威、公开透明的评价是学术评价的重要组成部分。20多年来,图书情报界、编辑出版界、大学、科研院所、学术团体、学术媒体和行政管理部门出于不同的评价目的,对学术图书进行过各种评价,积累了许多经验和教训,为较大规模地从内容质量上评价学术图书奠定了坚实的基础。学术“全评价”分析框架是借鉴、整合国内外评价理论和实践的产物,能够较为合理地解释各种学术评价现象和评价得失,具有一定的学理性、逻辑性、指导性和简明性。因此,在此框架下构建学术图书质量评价指标体系及其实施方案,也相应具有一定的合理性和创新性。学术图书质量评价要回归以学术共同体为主体的评价,引文等计量评价可以作为适当补充和校验,基于实践与时间(历史)的效应评价才是最有意义的评价。

(叶继元,南京大学信息管理学院特聘教授,博士生导师)

注释:

① 叶继元.学术图书、学术著作、学术专著概念辨析[J].中国图书馆学报,2016(1).  
② 叶继元.图书、学术图书与人文社科学术图书种数之考察[J].大学图书馆学报,2016(1).  
③ 2018年全国新闻出版业基本情况[EB/OL].(2019-08-29)[2020-08-08].http://www.cssn.cn/dubao/201908/t20190829\_4963940.shtml.  
④ 王兰敬,叶继元.中文人文社会科学学术图书评价的瓶颈因素及对策研究[J].图书与情报,2014(6):82-87.  
⑤ 冯平.评价论[M].北京:东方出版社,1995:156.  
⑥ 周可真.人文社会科学评价中的逻辑标准与事实标准[N].科学时报,2009-4-10.