

# 人文社会科学论文质量评估指标体系

Quality Evaluation Index System  
for Papers in Humanities & Social Sciences

## 实 施 方 案

中国人民大学人文社会科学学术成果评价研究中心

2010年12月

# 总 目 录

第 1 部分：指标体系的设计说明 .....	1
1.1 设计宗旨与目标 .....	1
1.2 设计理念与原则 .....	2
1.3 设计思路 .....	3
1.4 设计方法 .....	6
1.5 指标体系的特色 .....	7
第 2 部分：指标体系的基本设置 .....	10
2.1 论文分类 .....	10
2.1.1 评估对象界定 .....	10
2.1.2 分类方案 .....	11
2.2 同行评议指标体系 .....	12
2.2.1 指标的含义与评估内容 .....	13
2.2.2 指标权重分配 .....	15
2.3 文献计量指标体系 .....	16
2.3.1 指标的含义与评估内容 .....	16
2.3.2 指标权重分配 .....	17
第 3 部分：指标体系的组织实施 .....	19
3.1 同行评议的组织实施 .....	19
3.2 文献计量评估的实施 .....	27

3.3 综合评估.....	28
3.4 评估结果发布.....	29
第 4 部分：指标体系相关研究报告 .....	31
4.1 人文社会科学论文质量评估指标遴选文献调研报告 .....	32
4.1.1 文献调研概述.....	33
4.1.2 论文分类评估方案的调研.....	35
4.1.3 同行评议指标的调研 .....	41
4.1.4 文献计量指标的调研 .....	59
4.2 人文社会科学论文质量评估指标体系问卷调查报告 .....	70
4.2.1 问卷调查概述.....	71
4.2.2 论文分类评估方案 .....	73
4.2.3 论文评价指标的必要程度.....	76
4.2.4 论文评价指标的权重分配.....	84
附件：调查问卷.....	91
4.3 人文社会科学论文质量评估指标体系专家论证报告 .....	96
4.3.1 专家论证的研究概述 .....	97
4.3.2 指标体系的整体论证 .....	97
4.3.3 分类评估方案的论证 .....	97
4.3.4 基本指标设置的论证 .....	99
4.3.5 权重分配方案的论证 .....	103
4.3.6 体系实施方案的论证 .....	104
4.4 人文社会科学论文质量评估系统（2.0）使用指南.....	110
4.4.1 登陆.....	111
4.4.2 菜单管理.....	111
4.4.3 基本信息管理.....	112
4.4.4 评审管理.....	119
4.4.5 统计分析.....	129
4.4.6 信息查询.....	130

4.4.7 帮助功能.....	131
-----------------	-----

# 图 目 录

图 1	同行评议指标体系的论证路线.....	6
图 2	指标体系总体结构.....	10
图 3	论文权重分类设置方案.....	12
图 4	“21 分量表”示意.....	20
图 5	同行评议实施流程.....	21
图 6	文献计量评估实施流程.....	27
图 7	受访对象职业类型分布.....	73
图 8	全体受访者选择分类方案的调查.....	74
图 9	不同职业类型受访者分类方案的调查.....	76
图 10	受访者对“辅助指标”必要程度的看法.....	78
图 11	各学科“主要指标”和“辅助指标”权重分配对比.....	85
图 12	不同类型论文权重分配对比.....	85
图 13	“人文学科论文”和“社会科学论文”指标权重分配对比.....	86
图 14	“综述文章”和“研究论文”指标权重分配对比.....	87
图 15	“基础理论研究论文”和“应用实证研究论文”指标权重分类对比.....	87
图 16	各学科指标权重分配对比.....	88
图 17	系统登录界面.....	111
图 18	基本信息管理 - 课题立项界面.....	112
图 19	课题立项管理 - 修改课题立项界面.....	113

图 20 基本信息管理 - 报刊管理界面.....	114
图 21 基本信息管理 - 报刊管理界面.....	115
图 22 报刊管理 - 修改期刊界面.....	116
图 23 报刊管理 - 添加期刊信息界面.....	116
图 24 基本信息管理 - 权重管理界面.....	117
图 25 基本信息管理 - 二级学科管理界面.....	118
图 26 基本信息管理 - 地区管理界面.....	119
图 27 基本信息管理 - 专题信息管理界面.....	119
图 28 评审流程.....	120
图 29 分类管理.....	121
图 30 分类管理 - 论文分类.....	122
图 31 评分管理 - 论文评分.....	125
图 32 评分分布查询.....	126
图 33 明细查询.....	129
图 34 统计分析 - 统计报表查询.....	130
图 35 信息查询.....	131

# 表 目 录

表 1	评估对象确定的基本原则.....	11
表 2	同行评议指标的含义与评估内容 .....	13
表 3	同行评议主要指标与辅助指标的权重分配 (适用于所有论文) .....	15
表 4	“综述文章” 指标的权重分配 .....	15
表 5	各学科研究论文的指标权重分配 .....	16
表 6	文献计量指标的权重分配.....	18
表 7	论文分类表.....	22
表 9	同行评议论文定级百分比参考 .....	26
表 10	现有分类评估方案分析 .....	39
表 11	关于成果评估标准的文献分析 .....	41
表 12	根据文献调研入选的同行评议指标.....	56
表 13	根据文献调研弃用的指标 .....	57
表 14	文献调研结果与书报资料中心调研结果对比.....	58
表 15	国内现有文献计量评估指标体系对比 .....	60
表 16	文献计量指标遴选结果 .....	64
表 17	受访者的样本条件与数量 .....	72

表 18	受访者的学科分布.....	73
表 19	全体受访者对主要指标必要程度的看法 .....	77
表 20	“人文学科论文”与“社会科学论文”指标必要程度对比.....	79
表 21	“综述文章”和“研究论文”指标必要程度对比.....	81
表 22	“基础理论研究论文”和“应用实证研究论文”指标必要程度对比 ...	82
表 23	主要指标和辅助指标权重问卷调查结果 .....	84
表 24	复印报刊资料重要转载来源期刊(2014 年版)入选期刊数量.....	111

# 第 1 部分：指标体系<sup>1</sup>的设计说明

论文质量评估是针对一定时空范围内的特定需要进行的,而且受制于评估主体的认知能力和知识水平。由于人文社会科学成果质量的显现周期较长,不论在任何时空点上对其质量进行评估,都无法取得“完全准确”的结果。但这些并不能成为论文质量不可评或不可测的理由,否则科学研究活动就变得毫无意义。科学研究活动是一个“不间断的、从已知探索未知的过程”,因而对其成果的评估只能在“科学探究的终极目标”与“社会发展的现实需求”之间寻求平衡。毫无疑问,“相对科学可行”的评估对推动科学进步和社会发展,具有重要意义。

## 1.1 设计宗旨与目标

### 1.1.1 设计宗旨

汲取已有人文社会科学评估的理论成果精华和实践经验,立足当前实际应用,以“科学、谨慎、探索、实用”为原则,创新评估的理论方法,完善评估的指标体系,推动我国人文社会科学科研成果评估向科学化、规范化、应用化发展,促进学术研究的良性发展和科研管理水平的持续提升,推动我国人文社会科学事业的大繁荣、大发展。

### 1.1.2 设计目标

通过对人文社会科学论文质量评估所涉及的复杂变量因素进行深入细致地

---

<sup>1</sup> 本指标体系的研制,得到中国人民大学科研基金项目“人文社会科学学术论文质量评估的基本指标设置与实证研究”、ISTIC-THOMSON 科学计量学联合实验室开放基金项目“人文社科论文质量评估指标的探索研究”的资助,特此感谢。

研究分析，以及对现有指标体系进行剖析、改造和创新，针对当前评价现状与现实需求，把人文社会科学论文作为直接的评估对象，以“同行评议为主、文献计量为补充”，区分论文学科、体裁差异，分类评估、同类比较，强调创新、质量、服务导向，综合、平衡各种因素，构建一套相对科学可行的“人文社会科学学术论文质量评估指标体系”。

## 1.2 设计理念与原则

### 1.2.1 设计理念

#### (1) 质量导向

量化评估在当前的学术成果评估工作中占主导地位。但过于强调量化评估已导致学术成果的“数量激增”与“高水准成果稀缺”的尴尬反差。改善中国学术研究的环境，当务之急是要倡导“质量为本”的评估导向，克服单纯以成果数量评估学术水平和贡献的弊端<sup>2</sup>。本体系着重对论文内容质量进行评估，将直接反映质量的定性指标作为主要依据，将间接反映质量的定量指标作为辅助和补充。

#### (2) 引导创新

“创新性”是科学发展的灵魂，本体系强调评估对人文社会科学创新的引导作用：把论文的“学术创新程度”，作为评估论文质量的首要指标，占有较大比重；创新程度较低的论文，其最终评估等级也相应较低；将不具备创新性的论文直接评定为“较差”、“极差”级别<sup>3</sup>。

#### (3) 面向服务

本指标体系所形成的论文质量评估数据将大大弥补我国人文社会科学论文

---

<sup>2</sup> 郑文涛. 关于哲学社会科学评价的若干思考. 社会管理科学与评论[J]. 2008 (2) : 31-39.

<sup>3</sup> “综述文章”的创新性定义为“对已有成果做出新的概括、评析”，具体分析见本体系低 2 部分。

评估中缺乏同行评议数据的不足,同时也为科研管理、学科发展、个人学术研究、报刊编辑、期刊质量管理、同行研究等提供重要参考。以此为基础,将建立一个“开放的人文社会科学论文评估平台”,面向社会提供广泛的、专业化的评估信息服务。

### 1.2.2 设计原则

评估,是指主体依据一定标准对客体满足需要的能力和程度进行判断的活动,指标体系是判断的“具体标准”。本体系的设计主要遵循如下两项原则:

#### (1) 综合性原则

人们普遍认为,人文社会科学成果评估必须坚持直接指标与间接指标相结合、主观评估与客观评估相结合、学术价值与社会价值相结合等原则,否则极易造成评估的片面性。本指标体系的构建将从学术共同体的“共识性标准”出发,综合考虑上述各种因素。

#### (2) 平衡性原则

我们的理想是要建立一套科学的、完善的、可行的“评估指标体系”,但现实情况表明,指标体系的建立面临完备性与可行性、及时性与延迟性、精细化与高效率等一系列矛盾,以及影响人文社科成果评估的复杂因素,因而“面面俱到的完美体系”是不可能存在的。我们必须在“理论完美性”和“现实需求性”之间,对上述矛盾做出折中或妥协。

### 1.3 设计思路

综合平衡相关因素,我们对一个适合当前实际的、相对科学可行的“人文社会科学论文质量评估指标体系”做如下阐释。

### 1.3.1 关于评估对象：针对论文的“内容质量”而非论文的“形式载体”

本指标体系在本质上是对人文社会科学论文的内容质量而非形式载体进行评估，因此，指标的遴选将围绕“论文质量内涵”这个核心。质量是反映实体满足明确或隐含需要的能力的特征及特征总和<sup>4</sup>。它包含三个层次：1) 质量必须“满足需要”；2) 质量是满足主体需要的“能力属性”；3) 对质量的评估可以理解为对质的多少，即优劣程度的判断。因此，人文社会科学论文质量是指满足学术需要（提供新知识，揭示客观真理）和社会需要的能力属性，诸如创新、正确、有用、重要等；对质量的评估就是对论文的创新程度、正确程度、有用程度和重要程度等进行判断。

### 1.3.2 关于评估标准：以“真理标准”为主导，兼顾“社会价值标准”

“真理标准”和“价值标准”<sup>5</sup>是学术成果评估的两大主要标准。其中“真理标准”体现科学研究的“客观法则”，具有客观性和普遍性；“价值标准”则因主体需要的差异呈现出复杂性。首先，人文社会科学研究过程必须遵循人类认识活动的“客观法则”，因此其研究成果也必须接受“真理标准”的检验；同时，人文社会科学研究也是为人类社会发展服务的，因此也要接受“社会价值标准”的检验。本指标体系选择普遍适用于各学科和各类型论文的“真理标准”作为指标遴选的主导标准，兼顾“社会价值标准”，用以弥补“真理标准”难以反映的论文在“社会价值”方面的贡献。当然，“政治标准”是本体系坚持的前提条件，不符合政治标准的论文不纳入本体系进行评估，而由指定人员做单独评估。

<sup>4</sup> 国际标准化组织，ISO8402—1994《质量术语》。

<sup>5</sup> “真理标准”也称“学术标准”，含“创新标准”和“科学标准”；“价值标准”主要包括“学术价值”和“社会价值”两类标准；其中后者也称“非学术标准”或“应用标准”，含“效益标准”。

### 1.3.3 关于评估方法：同行评议为主，文献计量为补充

文献计量和同行评议是学术成果评估的两大主要方法。然而，单独采用任何一种评估方法，都有其局限性：由于引用行为的复杂性难以充分描述和统计，仅用文献计量法进行评估容易导致“重量轻质”、“重刊不重文”等问题；而评估主体的主观随意性较难控制，则是同行评议法的主要不足。因此，两种方法相结合的“综合评估”是学术评估领域的普遍共识。与自然科学研究相比，人文社会科学的研究具有更强的自由化和个性化等“小科学”特征<sup>6</sup>，成果转化利用的时间、范围、方式都比较模糊<sup>7</sup>。因此，我们认为应根据人文社会科学的研究规律，采用“同行评议为主，文献计量为补充”思路。

### 1.3.4 关于评估分类：区分差异、分类评估、同类比较

人文学科与社会科学之间，以及内部各细分学科之间，在研究方法和研究性质上均存在较大差异。因此，人文社会科学论文的评估不应该“一刀切”，必须针对论文之间的差异性，设置相应的指标和权重，进行分类评估、同类比较。然而，分类评估并非越细越好：较精细的分类评估有利于提高科学性，较粗略的分类评估有利于提高可行性。因此，本体系在确定分类评估方案时，充分权衡了“精细评估”与“粗略评估”的优势和不足，针对差异较为明显的论文特征（如综述文章与研究论文的差异、不同学科之间的差异等）设置单独分类，差异微小的特征不再进行分类评估。

### 1.3.5 关于评估实施：根据可有效获取数据的时间差异，发布多种评估结果

不同的论文评估指标所需的数据产生的时间不同：1) “学术创新程度”等定

<sup>6</sup> 朱少强. 论科学建制背景下的人文社会科学研究评价[J]. 评估与管理. 2008 (12): 44-48.

<sup>7</sup> 虞文, 周亚霆. 基于因素分析的社科成果转化评价指标浅探[J]. 科学学与科学技术管理:57-59.

性指标,在论文形成后可随时通过同行评议形成评估数据;2)“发表载体”指标的评估数据,需要在论文正式发表后才能获取;3)论文的 Web 下载频次、被转载次数等指标的评估数据,通常在论文发表半年至 1 年可有效形成;4)被引频次等文献计量指标,至少需要 3 年的数据积累才可计算出有效的数据。通常,延迟评估有利于得出更准确的结果,但及时评估更有利于人文社会科学学术成果出版、学术资源分配和对学术研究者的激励。现实表明,按年度形成评估结果是现阶段的迫切需求之一。为此,我们重点针对论文发表当年可形成有效数据的一系列指标,设计每年定期对前一年发表的论文质量进行评估的指标体系实施方案。

## 1.4 设计方法

### 1.4.1 论文质量评估指标遴选的原则

- (1) **必要性** (认同范围): 理论上和实践上得到学者和评估者的广泛认同。
- (2) **有效性**: 与论文质量正相关,能有效反映论文质量。
- (3) **相对独立性**: 在逻辑上与其他指标不重合,其他指标不可替代。
- (4) **可操作性**: 指标内涵清晰、易于评估把握。

### 1.4.2 同行评议指标体系的设计方法

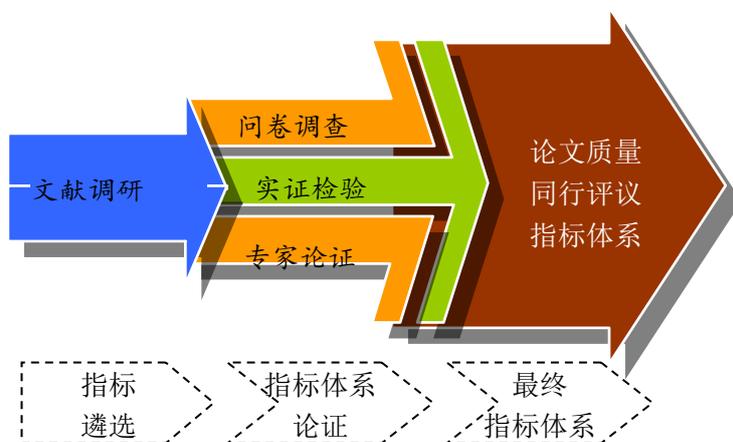


图1 同行评议指标体系的论证路线

**(1) 文献调研:**汇集、梳理、归纳、总结出人文社会科学学术论文质量评估的基本指标 (见 4.1 部分)。

**(2) 问卷调查:**针对基本指标的必要性、适用性和重要性 (即权重), 向具有代表性的学术报刊编辑和具有较强学术影响力的学科专家进行问卷调查。指标的权重分配采用层次分析法, 通过构建指标“两两比较矩阵”确定 (见 4.2 部分)。

**(3) 专家论证:**邀请同行专家对指标体系的创新性、科学性和可操作性等进行总体论证 (见 4.3 部分)。

**(4) 实证检验:**通过评估 1.5 ~ 2 万篇以上的学术论文样本, 对指标体系进行分学科、大规模的实际应用和验证。

### 1.4.3 文献计量评估指标体系的设计方法

**1) 文献调研:**汇集、梳理、归纳、总结出可用于人文社会科学单篇学术论文质量评估的文献计量指标 (见 4.1 部分)。

**2) 推理分析:**分析所遴选出的文献计量指标, 对于评估人文社会科学单篇学术论文是否有效、可操作, 并对其进行推理分析, 最终确定出可对论文同行评议进行补充的文献计量评估指标。

## 1.5 指标体系的特色

### 1.5.1 以学术论文为直接评估对象

本体系以学术论文为直接评估对象, 有效规避了“以刊评文”带来的种种弊

端。同行评议和文献计量指标体系的设计,都强调对论文内容质量的评估,把“有效反映论文质量”作为遴选指标的必要条件,主要表现在:以直接反映论文内容质量的定性指标为主,如学术创新程度、论证完备程度、社会价值、难易程度;以间接反映论文质量的指标为辅,如课题立项、发表载体;以定量性较强的文献计量指标为补充,如 Web 下载频次、被他引频次、被转载次数。

### 1.5.2 坚持同行评议的主体地位

当前,我国各学科论文水平评估中,文献计量评估占据主导地位。然而实践表明,同行评议更符合人文社会科学的研究特点和规律,更有利于及时有效地发现优秀成果;文献计量评估则有利于弥补同行评议主观随意性的不足。本体系坚持同行评议的主体地位,兼顾文献计量评估的优势,既顺应了学术评估领域的客观需求,又有利于弥补当前我国人文社会科学论文评估的不足。

### 1.5.3 指标通用、分类设置权重、分步评估

同行评议指标是本体系的设计重点,为使该指标体系既能较为合理地体现不同学科和类型论文的特质,又具有较强的可操作性,我们采用了“指标通用、分类设置权重、分步评估”的设计思路:

(1) 从学术共同体的“共识性评估标准”中,提取出普适性较高的指标,设置通用的指标,用于评估所有论文。

(2) 针对论述体裁和学科领域的差异,分类形成不同的指标权重分配方案。

(3) 设置三个评估实施环节,首先通过“初评”评估论文的学术性质和基本水平,筛选掉非学术性论文;之后通过“复评”判断论文基本的“优劣等级”,

筛选掉质量较差的论文，并对质量一般以上的学术论文进行评估；最后通过“终评”对质量一般以上的学术论文进行再次评估。

## 第 2 部分：指标体系的基本设置

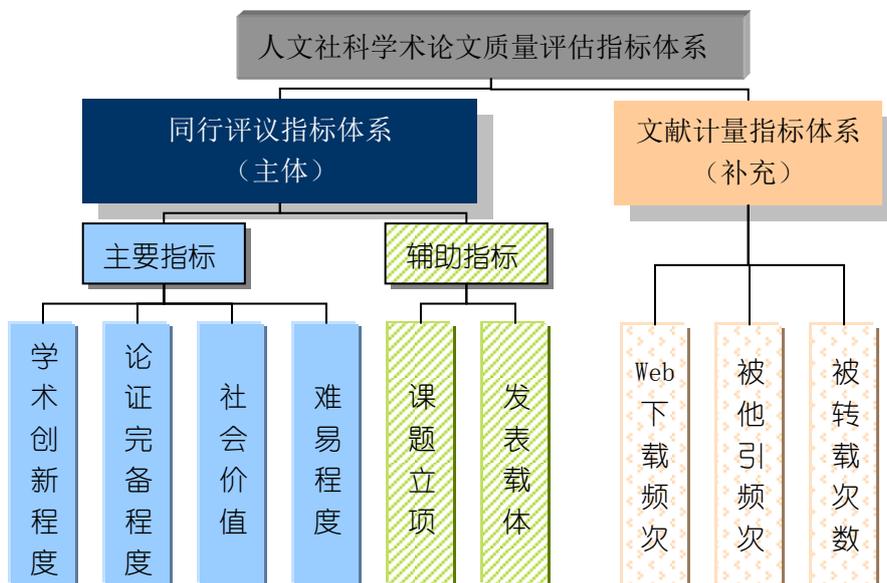


图 2 指标体系总体结构

如图 2 所示，指标体系分为同行评议指标体系和文献计量指标体系两大部分。其中，前者占主体地位，后者作为必要补充。同行评议指标体系中，又分为主要指标和辅助指标，最终将形成直接的、及时的论文评估结果；文献计量指标体系中 3 个指标形成的评估结果将起补充作用。

### 2.1 论文分类

#### 2.1.1 评估对象界定

本指标体系的评估对象是“所有正式发表的人文社会科学学术论文”<sup>8</sup>。从论述体裁上可分为两大类：

**(1) 研究论文：**是指研究者在探讨人文社会科学某一学科对象过程中形成的、记录并阐述研究的创新见解和知识、并按照特定范式撰写的书面文件，可用

<sup>8</sup>本体系重点阐述“已正式发表论文”的评估实施方案，未来将逐步扩展到“未发表论文”，以及其他类型的人文社会科学研究成果的质量评估。

于学术会议上宣读、交流、讨论或在学术刊物上发表。研究论文的基本特征为：学术性、创新性、理论性、规范性。

**(2) 综述文章：**是指对人文社会科学某学科领域或专题在某段时期的研究状况进行系统地查阅、整理、总结和分析，并做出综合性描述或评议的文章。

人文社会科学学科范围，暂以《授予博士、硕士学位和培养研究生的学科、专业目录》(1997)所确定的人文社会科学学科为基本依据。

对于译文、学术动态，领导讲话、政策解析、标准技术规范解析，书评、理论学习、社会实践札记，时事、政治、思想评论，文学评论、艺术评论等文章，在确定是否为评估对象时，应参照下述原则（见表1）。

表1 评估对象确定的基本原则

论文特征	作为评估对象	不作为评估对象
目的	以探讨一般规律为目的	以介绍、宣传或推荐为目的
读者对象	以学者为读者对象	以社会大众为读者对象
论文内容	有明确研究过程并有所创新的文章	未经过研究过程、无创新的文章

【说明】对译文质量的评估，是对原文作者（而非译者）的学术研究成果进行评估。

### 2.1.2 分类方案

通过文献调研（见4.1部分）和问卷调查（见4.2部分）可知：1) 按学科进行分类评估、同类比较，得到学术共同体和评价实践者的广泛认同；2) 利用同行评议法评估不同学科论文时，差异主要体现在权重设置上，而非指标设置上；3) 综述文章与研究论文的指标权重分配，存在明显差异；4) 各一级学科论文的指标权重分配，存在一定差异。基于上述结论，本指标体系的论文分类方案如下：

**(1) 同行评议指标体系采用“指标通用、分类设置权重、同类比较”的方案**

- 1) 对人文社会科学所有论文，设置通用的评估指标。
- 2) 根据论文的论述体裁和所在学科，设置不同的指标权重分配方案（见图3）。
- 3) 首先分为“研究论文”和“综述文章”（所有学科的“综述文章”设置统一的指标权重）；“研究论文”按所在一级学科，设置不同的权重分配方案。
- 3) 分析评估分数时，对相同学科的论文进行同类比较。

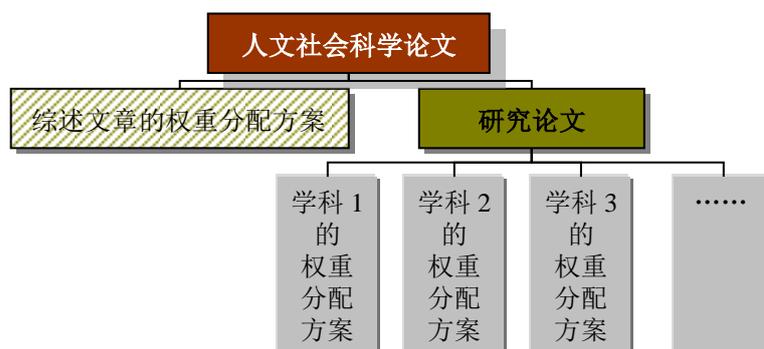


图3 论文权重分类设置方案

**(2) 文献计量指标体系采用“指标和权重通用，同类比较”的分类评估方案**

- 1) 针对所有人文社会科学论文，设置通用的评估指标和权重分配方案。
- 2) 分析评估分数时，对相同学科的论文进行同类比较，学科分类规则与同行评议部分相同。

**【说明】**

- 学科分类暂定以教育部《授予博士、硕士学位和培养研究生的学科、专业目录》(1997)所确定的人文社会科学学科为参照；评估分析时，可根据实际评估需要(如论文的规模)，对学科分类进行适当调整。
- “跨学科论文”，一般是指论文的研究内容涉及一个以上的一级学科。分为两类：
  - “跨学科研究论文”：围绕某一主要学科的理论、方法或问题开展研究，但借鉴其他学科方法和理论。此类论文在该主要学科中进行评估。
  - “交叉学科研究论文”：研究同时涉及一个以上学科的理论、方法或问题，学科归属不明晰。此类论文同时在不同学科中进行独立评估。

**2.2 同行评议指标体系**

同行评议指标体系是同行专家运用专业知识和判断能力,对论文质量进行直接评估的依据。

### 2.2.1 指标的含义与评估内容

表 2 同行评议指标的含义与评估内容<sup>9</sup>

类别	评估指标	指标内涵	评估内容
主要指标	学术创新程度	衡量论文提供的新知识对学术发展的促进程度。	<p><b>以下内容对学术发展的促进程度：</b></p> <ul style="list-style-type: none"> <li>➢ 提出新的（或修正完善已有的）学说、理论、观点、问题、阐释等</li> <li>➢ 提出新的（或改进运用已有的）方法、视角等</li> <li>➢ 发现新的资料、史料、证据、数据等</li> <li>➢ 对已有成果做出新的概括、评析（仅指综述文章）</li> </ul>
	论证完备程度	衡量论文的研究规范程度和严谨程度。	<p><b>研究方法有效性：</b></p> <ul style="list-style-type: none"> <li>➢ 研究方法科学性</li> <li>➢ 研究方法适当性（对于研究问题）</li> </ul> <p><b>论据可靠性：</b></p> <ul style="list-style-type: none"> <li>➢ 资料占有全面性</li> <li>➢ 资料来源真实性</li> <li>➢ 资料引证规范性</li> </ul>

<sup>9</sup> 通过文献调研（4.1 部分）和问卷调查（4.2 部分），抽取“学术共同体”较为普遍认同的指标以及评估内容，纳入本体系。

类别	评估指标	指标内涵	评估内容
			<b>论证逻辑性:</b> <ul style="list-style-type: none"> <li>➢ 理论前提科学性</li> <li>➢ 概念使用准确性</li> <li>➢ 论证过程系统性</li> <li>➢ 逻辑推理严密性</li> </ul>
	社会价值	衡量论文对社会发展进步可能产生的推动作用的大小。	<ul style="list-style-type: none"> <li>➢ 对解决经济、政治、社会建设中问题的推动作用</li> <li>➢ 对思想道德文化建设的促进作用</li> </ul>
	难易程度	衡量论文研究投入劳动的多少。	<b>论题复杂度:</b> <ul style="list-style-type: none"> <li>➢ 理论难点的多少</li> <li>➢ 实证研究的难度</li> </ul> <b>资料难度:</b> <ul style="list-style-type: none"> <li>➢ 资料搜集难度</li> <li>➢ 资料处理难度</li> </ul>
辅助指标	课题立项	论文来源的课题立项的情况。	<ul style="list-style-type: none"> <li>➢ 国家级 (21分)</li> <li>➢ 省部级 (14分)</li> <li>➢ 其它立项 (8分)</li> <li>➢ 无立项 (1分)</li> </ul>
	发表载体	论文发表载体的学术影响力。	<ul style="list-style-type: none"> <li>➢ 核心报刊 (21分)</li> <li>➢ 非核心报刊 (11分)</li> </ul>

【说明】

- “辅助指标”也体现了论文形成和发表过程中同行专家的评估意见，一般都与论文质量关系密切，有必要作为评估的辅助指标，因此纳入本体系（具体分析见 4.1 部分）。
- 课题立项是指资助单位所处的层级。国际、横向和商业项目，将由评委根据实际情况，确定相应等级。由于国际项目较少，不再单独设置细分等级，根据国际资助机构层级，归入国家级、省部级或其它立项；高等院校研究基金资助项目，统一归入“其它立项”。
- 发表载体是指论文发表当年被认定的报刊分级结果。其中“核心报刊”是指以下“四大评估体系”中任何一个体系所认定的报刊：
  - 北京大学《核心期刊要目总览》（2008）
  - 南京大学 CSSCI 来源期刊（2010-2011）
  - 中国社科院《中国人文社会科学核心期刊要览》（2008）
  - 复印报刊资料重要转载来源期刊（2009）（见 4.5 部分）

由于刊载学术论文的报纸较少，不再单独设置等级。除 4.5 部分列表中的报纸外，其它报纸的等级可根据行业影响情况由评委确定。

## 2.2.2 指标权重分配<sup>10</sup>

### (1) 主要指标和辅助指标的权重分配

表 3 同行评议主要指标与辅助指标的权重分配（适用于所有论文）

类别	权重
主要指标	85%
辅助指标	15%

### (2) 综述文章的权重分配

表 4 “综述文章”指标的权重分配

类别	指标	综述文章
主要指标 (85%)	学术创新程度	29%
	论证完备程度	29%
	社会价值	27%
	难易程度	15%

<sup>10</sup> 本体系的指标权重分配方案，根据问卷调查结果（见 4.2 部分）及专家论证结果（见 4.3 部分）论证得出。

辅助指标 (15%)	课题立项	54%
	发表载体	46%

### (3) 各学科研究论文的指标权重分配

表 5 各学科研究论文的指标权重分配

类别	指标	哲学	理论经济学	应用经济学	法学	政治学	社会学	民族学	马克思主义理论	教育学	心理学	体育学	中国语言文学	外国语言文学	新闻传播学	艺术学	历史学	地理学	管理科学与工程	工商管理	农林经济管理	公共管理	图情档
		主要 指标 85%	学术创新程度	38%	36%	41%	32%	33%	36%	36%	34%	37%	39%	34%	41%	41%	38%	46%	34%	34%	29%	29%	29%
论证完备程度	28%		27%	24%	24%	24%	25%	25%	21%	24%	21%	25%	24%	24%	23%	24%	25%	25%	29%	29%	29%	29%	25%
社会价值	18%		21%	22%	26%	28%	23%	23%	31%	21%	29%	24%	18%	18%	28%	18%	26%	26%	30%	30%	30%	30%	25%
难易程度	16%		16%	13%	18%	15%	16%	16%	14%	18%	11%	17%	17%	17%	11%	12%	15%	15%	12%	12%	12%	12%	11%
辅助 指标 15%	课题立项	33%	53%	40%	40%	47%	53%	53%	40%	50%	40%	40%	53%	53%	47%	33%	50%	50%	47%	47%	47%	47%	47%
	发表载体	67%	47%	60%	60%	53%	47%	47%	60%	50%	60%	60%	47%	47%	53%	67%	50%	50%	53%	53%	53%	53%	53%

## 2.3 文献计量指标体系

### 2.3.1 指标的含义与评估内容

通过对现有文献计量评估指标的汇总和分析（见 4.1 部分），按照“可针对单篇论文进行统计分析、可有效反映论文学术影响力”的遴选原则，确定以下 3 个文献计量指标。

### (1) 论文 Web 下载频次 (月均<sup>11</sup>)

指论文发表后一段时间内, 每月平均被网络下载的频次; 下载频次越高, 论文的影响力越大。计算公式如下:

$$\text{Web 下载频次} = \frac{\text{论文自上网后被下载的频次之和}}{\text{论文统计时间}-\text{论文上网时间(月)}}$$

#### 【说明】

- 为防止人为增加下载频次, “相同 IP” 在 24 小时内下载同一篇论文, 只算作 1 次。
- 每年定期对前 1 年和 3 年前发表的论文<sup>12</sup>的 “Web 下载频次” 进行分别统计。

### (2) 论文被转载次数

指论文发表后一段时间内, 被《新华文摘》、中国人民大学复印报刊资料、《中国社会科学文摘》、《高等学校文科学术文摘》转载的次数。被转载次数越多, 表示论文的影响力越大。

#### 【说明】

- 分数确定为 5 级: 4 大文摘中, 4 次转载评分为 1, 3 次转载评分为 0.8, 2 次转载评分为 0.6, 1 次转载评分为 0.4, 0 次转载评分为 0。
- 通常, 论文发表后 1 年以内即可确定是否被转载, 因此本指标体系每年定期对前 1 年新发表论文<sup>13</sup>的 “转载次数” 进行统计。

### (3) 论文被他引频次

指论文发表后一段时间内, 被他人引用的总频次; 被引用的频次越高, 论文的影响力越大。

#### 【说明】

- 施引论文的 “所有作者” 与被引论文的 “所有作者” 不重叠时算作 “他引”。
- 本指标体系每年定期对 3 年前发表论文<sup>14</sup>的 “被他引频次” 进行统计。

## 2.3.2 指标权重分配

上述指标分别从不同角度反映了论文被传播、被使用和被认可的情况, 各指

<sup>11</sup> 设置 “月均” 下载频次, 是为了防止论文在一段时间中上网时间不同所导致的不公平。比如统计某年 1 月-12 月的论文下载频次, 1 月发表的论文因为上网的时间长, 所以被下载的机会更大。

<sup>12</sup> 比如 2010 年 1 月, 集中统计 2009 年 1 月-12 月发表的论文, 以及 2007 年 1 月-12 月发表的论文。

<sup>13</sup> 比如 2010 年 1 月, 集中统计 2009 年 1 月-12 月发表的论文的 “被转载次数”。

<sup>14</sup> 比如 2010 年 1 月, 集中统计 2007 年 1 月-12 月发表的论文。南京大学苏新宁教授研究认为, 72% 的人文社会科学期刊论文的引用峰值, 出现在论文发表后的 3 年内 (《2009 年中国人文社会科学期刊学术影响力报告》[M], 中国社会科学出版社: 2009: 13)。

标之间是并列关系。通过课题组反复讨论和专家论证最终确定权重值如下：

表 6 文献计量指标的权重分配

指标	权重	
	1 年数据	3 年数据
论文 Web 下载频次	0.4	0.2
论文被转载次数	0.6	0.3
论文被他引频次	—	0.5

【说明】只有相同时间区间内的文献计量指标的评估数据可进行加权求和。若某一文献计量指标的数据无法获取，该项指标的权重将按其它两个指标的权重比例，分配到其它两个指标中。

## 第 3 部分：指标体系的组织实施

### 3.1 同行评议的组织实施

#### 3.1.1 评估主体

在各学科领域内，选择具有良好学术道德和信誉、治学严谨客观公正、热心人文社会科学事业、责任心强的同行专家，组建评估团队。分为三个小组：

##### (1) 初评组

初评组负责对所有论文进行初步评估，鉴别论文的学术性，并对学术论文进行分类。初评评委应熟悉人文社会科学学术论文的基本特征和分类，能较好地把握论文的学术特征。

##### (2) 复评组

复评组负责对论文进行复评工作。复评评委应大量、持续地阅读所负责学科领域的学术成果，熟悉该领域的学术创新情况和发展动态，能较好地把握评估尺度，区分论文的优劣等级。

##### (3) 终评组

终评组负责对进入终评的论文进行评估。终评评委应具有较高的专业知识水平（拥有学科领域代表性成果）和实践经验、敏锐的洞察力和较强的判断能力，熟悉被评估内容及国内外相关领域的发展状况，并在此专业领域内具有一定权威性，能较好地把握评价标准。

【说明】

- 所有评委在正式开展评估工作之前，均须进行相应培训和“试评估”。
- 系统自动对评委的评估活动进行记录和分析，作为改进评估行为和更新专家库的依据。

### 3.1.2 评估方法与工具

#### (1) 21 分量表

同行评议过程是对主观印象的表达。指标体系的打分表，相当于一个“心理量表”<sup>15</sup>。该表须解决的核心问题是：较为准确地获得评估主体印象；尽量减少主观印象表述和量化过程中的信息损失。本指标体系的同行评议打分操作，选择“等距量表”。5 分量表、7 分量表、9 分量表或 10 分量表是常用的“等距量表”。然而，心理学研究表明<sup>16</sup>，人的短时记忆从 4-9 个信息单位不等，评分过程也符合短时记忆容量的规律。因此，我们选择了两个 5 分量表嵌套的“21 分量表”<sup>17</sup>。要求评委在打分时，先判断论文水平处于五个基本等级中的哪一级，再左右微调论文的最终得分。这种方法避免了因分级过少而强行做出选择时造成的评估信息丢失，或因分级过多导致超出评委判断能力造成的失误。

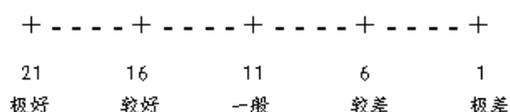


图 4 “21 分量表” 示意

#### (2) 评估软件

评估软件的主要功能包括：论文信息管理，指标体系管理，评估分数录入、查看、修改、计算合成，评分排序统计等功能（参见 4.4 部分）。

### 3.1.3 评估流程

按月度循环组织同行专家按照“同行评议指标体系”对论文质量进行初评、

<sup>15</sup> 百度百科“心理量表”是用来量化评估者观察中所得印象的一种测量工具。<http://baike.baidu.com/view/1639501.htm>。

<sup>16</sup> 米勒（G.Miller）于 1956 年提出，人的短时记忆容量为 7+2 个信息单位。这一短时记忆容量又叫短时记忆广度。1974 年西蒙（H.A.Simon）提出，人的实际工作记忆广度小于 7，只有 4 或 5 个信息单位。

<sup>17</sup> 卜卫，周海宏等. 社会科学成果价值评估[M].北京：社会科学文献出版社. 1999：157。

复评和终评；每篇论文共由 **5-7 名评委** 完成全流程的评估；每年定期对数据进行汇总和统计分析，并形成评估成果（见图 5）。

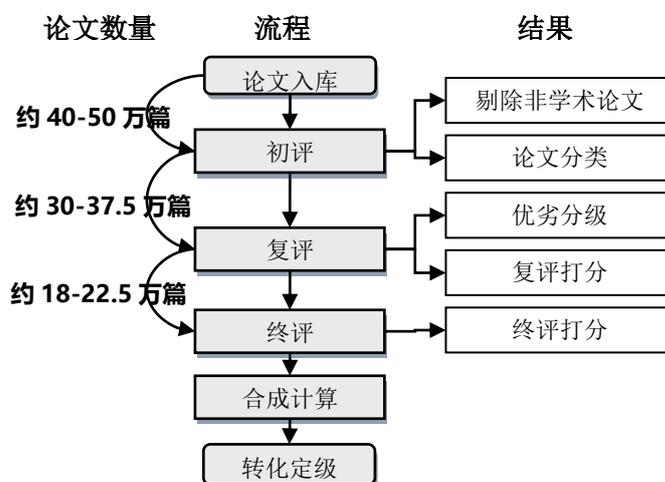


图 5 同行评议实施流程<sup>18</sup>

### 3.1.4 实施细则

#### (1) 论文入库

入库内容为论文的基本信息<sup>19</sup>：论文标题、作者、作者单位、关键词、摘要、原发报刊名称，刊期和起始页码等（详见 4.4 部分）。

【说明】论文的“发表载体”，由评分软件从报刊基本信息中自动提取。发表载体分级标准参见 2.2.1 部分。

#### (2) 初评

论文初评既是对论文基本学术性的鉴别过程，也是对论文的分类过程。

1) 初评环节将筛选掉非学术性的论文<sup>20</sup>，非学术性论文不进行分类，也不进入下一个评估环节；筛选标准参照本体系 2.1.1 部分对评估对象的说明。

2) 鉴别为学术论文后，初评人员对论文的所在学科、专题和论述体裁(如表

<sup>18</sup> 初步估计，我国（不含港澳台地区）学术性较强的人文社会科学报刊近 4000 种，每年刊载论文约 40-50 万篇。

<sup>19</sup> 本实施方案，暂时采用线下阅读论文、在线评分的方式，未来将实现全流程在线评审。

<sup>20</sup> 不符合“政治标准”的论文不纳入本体系的评估范围，另行评价。

7 所示)对论文进行多维分类。

表 7 论文分类表

论文属性	选项
论文所在学科 <sup>21</sup>	-哲学 <input type="checkbox"/> 中国哲学 <input type="checkbox"/> 西方哲学 <input type="checkbox"/> 逻辑学..... +法学 +体育学 +政治学 +艺术学 +新闻传播学 .....
论文所在专题 <sup>22</sup>	<input type="checkbox"/> A1 <input type="checkbox"/> A2 .....
是否为综述文章 <sup>23</sup>	<input type="radio"/> 是 <input type="radio"/> 否

**【说明】**

- “论文所在学科”划分到二级学科，可多选。“跨学科研究论文”在其所属的主学科评估；“交叉学科研究论文”由不同学科的专家同时独立评估。
- 论文分类功能参照软件功能设计报告中的“评审界面”（见 4.4 部分）。

**(3) 复评**

1) 核对分类：在复评之前，复评评委应对分类进行核对，可对所属分类进行补充和完善。若修订原有分类，需与初评人员沟通确认。

2) 优劣分级：每个学科的复评评委根据表 8 要求，在评分软件里对论文进行独立评估。复评将把论文分为两部分：

- **质量一般及以上的论文**：即表 8 中所示 A、B、C 三级论文，比例约为进入复评论文的 60-70%<sup>24</sup>。由复评评委根据表 8 在软件里评分（在 A、B、C 三级内进行评分）后，进入终评。
- **质量明显较差的论文**：即表 8 中所示的 D、E 两级论文，比例约为进入

<sup>21</sup> 暂以教育部《授予博士、硕士学位和培养研究生的学科、专业目录》(1997 年颁布)所确定的人文社会科学学科分类为准。

<sup>22</sup> 参照《书报资料中心复印报刊资料分类指南》，2010 年版。

<sup>23</sup> 综述文章的定义，参见本体系 2.1.3 部分。

<sup>24</sup> 约 18-22.5 万篇。

复评论文的 30%-40%<sup>25</sup>。这部分论文不进行终评，复评不按指标进行详细评分，只评出论文的总体等级，以复评等级作为最终评估结果。

**【说明】**

- 若同一学科的多名复评评委对同一篇论文“是否进入终评”的判断不一致，则该篇论文算作“**进入终评**”；若评委一致认为该篇论文“不进入终评”，但对该篇论文的等级判断不一致，则该篇论文算作“**D级**”。
- “课题立项”指标的评估标准较为明确（参见 2.2.1 部分表 1），复评评委确定立项等级后，即作为该指标的最后结果，终评评委不需要再次评估。

#### **(4) 终评**

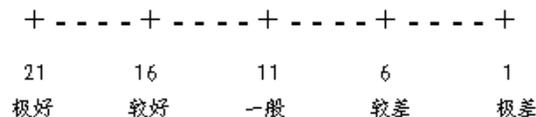
终评评委根据指标体系（见表 8）阅读论文，并在评分软件里对进入终评的论文进行独立评分（在 A、B、C 三级内进行评分）。

**【说明】**进入终评的论文，将综合所有复评评委和终评评委的评分，确定论文的最终分数和等级（具体操作见本体系 3.2.3）。

---

<sup>25</sup> 约 12-15 万篇。

【打分说明】请评委依次根据以下指标评分； 打分时请根据对指标各级得分的描述，选择 1-21 之间的数字（如右图所示），填入下表最右侧一列。如果在评估“学术创新程度”指标时，出现表中 D、E 的情况时，可把“论文等级”直接定为“D 较差”或“E 极差”，不再进行其它指标的打分。



指标解释与评估依据

•**学术创新程度**：衡量论文提供的新知识对学术发展的促进程度。

-评估内容包括：

- 1) 提出新的 (或修正完善已有的) 学说、理论、观点、问题、阐释等；
- 2) 提出新的 (或改进运用已有的) 方法、视角等；
- 3) 发现新的资料、史料、证据、数据等；
- 4) 对已有成果做出新的概括、评析 (仅针对综述文章)。

•**论证完备程度**：衡量论文的研究规范程度和严谨程度。

-评估内容包括：

- 1) 研究方法有效性；
- 2) 论据可靠性；
- 3) 论证逻辑性。

•**社会价值**：衡量论文对社会发展进步可能产生的推动作用的大小。

-评估内容包括：

- 1) 对解决社会和经济发展中问题的推动程度；
- 2) 对人类思想进步的促进程度。

•**难易程度**：衡量论文研究投入劳动的多少。

-评估内容包括：

- 1) 论题的复杂程度 (理论难点或实证研究难度)；
- 2) 资料搜集处理难度。

表 9 人文社会科学论文质量评估指标打分表

指标	21 分 (A 极好)	16 分 (B 较好)	11 分 (C 一般)	6 分 (D 较差)	1 分 (E 极差)	得分
学术创新程度	全面创新, 对学术发展有显著贡献。	有较大创新, 对学术发展有较大贡献。	有一定新意, 对学术研究有启发和参考意义。	选题陈旧, 无新意。	选题无意义, 内容雷同。	
论证完备程度	研究方法科学先进。	研究方法较为科学。	研究方法基本可行。	研究方法不适当。	研究方法无效。	
	论据全面准确; 引证清晰规范。	主要论据充实正确, 细节部分略有欠缺; 引证清晰规范。	主要论据基本正确, 但欠充实; 引证基本规范。	重要论据有遗漏; 引证不规范。	论据严重不足, 存在错误; 引证不规范。	
	理论前提科学, 概念使用准确, 推理严密系统, 分析深入透彻。	论证说服力较强, 分析较深入。	论证基本清晰, 分析有一定深度。	论证缺乏说服力, 分析浅显。	论证混乱, 缺乏分析。	
社会价值	对社会基本问题提出了切实可行的解决思路; 或对人类思想进步具有较大的推动作用。	对社会基本问题提出了较好的解决思路; 或对人类思想进步具有一定的推动作用。	对促进社会基本问题的解决或人类思想进步, 具有启发或参考意义。	基本无社会价值。	无社会价值。	
难易程度	论题研究难度大。	论题研究难度较大。	论题有一定研究难度	论题较为简单。	论题简单。	
	资料搜集处理困难。	资料搜集处理较困难。	资料搜集处理具有一定难度。	资料搜集处理难度较小。	资料搜集处理无难度。	

请判断您对该篇论文的总体印象：A 极好 B 较好 C 一般 D 较差 E 极差

## (5) 分数合成计算

### 1) 未进入终评论文的最终分数

由评分软件根据复评论文的等级，统一给出最终分数。

### 2) 进入终评论文的最终分数

#### ①原始得分的合成计算

以管理学研究论文为例，设第  $i$  个评委评估第  $j$  篇论文的总分为  $S_{ji}$ ：

$$S_{ji} = 0.85 \times (0.29 \times \text{学术创新程度} + 0.29 \times \text{论证完备程度} + 0.30 \times \text{社会价值} + 0.12 \times \text{难易程度}) + 0.15 \times (0.47 \times \text{课题立项} + 0.53 \times \text{发表载体})$$

#### ②得分数据的修正处理

为了减少不同评委打分尺度的主观差异性，需要对每位评委的所有评分结果进行修正。修正方法如下<sup>26</sup>：

设某一论文集中共有  $n$  篇论文参评，复评和终评共有  $m$  位评委，第  $i$  个评委给第  $j$  篇论文的评分为  $S_{ji}$ ，则：

第  $i$  个评委对  $n$  篇论文评分的平均分：

$$\bar{S}_i = \frac{\sum_{j=1}^n S_{ji}}{n} \quad (i=1,2,\dots,m; j=1,2,\dots,n) \quad (\text{公式 1})$$

$m$  个评委打分的区间平均值：

$$k = \frac{\sum_{i=1}^m \min(S_{1i}, S_{2i}, \dots, S_{ni})}{m} + \frac{\sum_{i=1}^m \max(S_{1i}, S_{2i}, \dots, S_{ni})}{m} \quad (i=1,2,\dots,m; j=1,2,\dots,n) \quad (\text{公式 2})$$

那么，第  $i$  个评委评估第  $j$  篇论文修正后的值为  $U_i$ ：

<sup>26</sup> 参考：程淑，桂林，冀航. 主观评分的归一化算法及误差分析[J]. 高等函授学报(自然科学版). 2007(10): 28-30.

$$U_i = \frac{kS_{ji}}{S_i} \quad (i=1,2\dots m; j=1,2\dots n) \quad (\text{公式 3})$$

### ③单篇论文的最终分数

对  $m$  位评委的修正分数进行平均后, 得到第  $j$  篇论文的最终分数为  $\bar{U}_j$  :

$$\bar{U}_j = \frac{\sum_{i=1}^m U_i}{m} \quad (i=1,2\dots m; j=1,2\dots n) \quad (\text{公式 4})$$

## (6) 转化定级

转化定级是指对**进入终评**的论文, 将其指标评分结果**转化为**论文质量**级别**的过程。论文质量级别实际上是表示单篇论文质量在某论文集合中(如 12 个月内、相同学科的论文) 所处的百分比位置。

1) **论文得分排序**: 对一定时期内相同学科所有论文的最终分数(修正平均后) 进行排序。“**跨学科研究论文**” 在所归属的主要学科内进行排序, 形成一个评估结果; “**交叉学科研究论文**”, 在不同学科内分别进行排序, 形成多个评估结果, 其中**最高得分**作为该论文的综合得分。

【说明】“跨学科研究论文” 和 “交叉学科研究论文” 定义见 2.1.1 部分。

2) 根据排序结果, 确定论文级别的划分, 定级标准为 “**占进入终评论文的百分比**”。本体系测定进入终评环节的各级论文的比例为:

表 9 同行评议论文定级百分比参考

百分比 级别	占进入终评论文的 百分比 (约 18-22.5 万)	占所有学术论文的 百分比(约 30-37.5 万)	占所有论文的百分 比 (约 40-50 万)
A “极好论文”	约前 10% (约 1.8-2.25 万)	约前 6%	约前 4.5%

B “较好论文”	约前 11%-30% (约 3.7-6 万)	约前 7%-16%	约前 4.6%-12%
C “一般论文”	剩 余 70% ( 约 12.7-14 万)	约前 17%-40%	约前 13%-30%

【说明】上表中的百分比为“预估值”。由于每年论文数量、质量存在波动，实际操作时，可参考上述标准进行微调。

## 3.2 文献计量评估的实施

### 3.2.1 评估工具

文献计量评估的分数，依托评分软件进行辅助计算。评估软件的主要功能包括：文献计量数据统计，分数的录入、查看、修改、计算合成、排序等<sup>27</sup>。

### 3.2.2 评估流程

根据如下流程，每年定期对前 1 年发表的和 3 年前发表的论文的文献计量指标得分，分别进行统计分析。

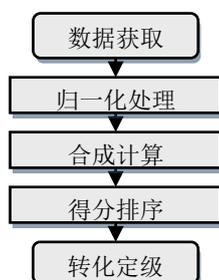


图 6 文献计量评估实施流程

### 3.2.3 实施细则

#### (1) 归一化处理

由于文献计量指标得分的量纲不同，需要进行归一化处理。将所有指标得分的值域映射为 (0,1) 区间的数字。归一化方法是,将每一指标中的最高者设置为

<sup>27</sup> 评估系统的文献计量指标统计功能，与同行评议功能类似。

1, 其他数值依据相应比例获取归一化后的数值。

设某一论文集合中的第  $i$  篇论文的某一指标值为  $N_i$ , 归一后的值为  $U_i$ , 则:

$$U_i = \frac{N_i}{\text{Max}\{N_i, |_{i=1,2,\dots}\}}$$

其中,  $\text{Max}\{N_i, |_{i=1,2,\dots}\}$  是指这一指标最大的一个评分数值。

## (2) 合成计算

论文文献计量评估的分数计算公式如下 (权重参见本报告 2.3.2):

### 1) 对 1 年的评估数据的合成计算

文献计量指标分数 =  $0.4 \times$  “Web 下载频次归一化值” +  $0.6 \times$  被转载次数

### 2) 对 3 年的评估数据的合成计算

文献计量指标分数 =  $0.2 \times$  “Web 下载频次归一化值” +  $0.5 \times$  “被他引频次归一化值” +  $0.3 \times$  被转载次数

## (3) 论文定级

每年定期对前 1 年发表的论文和 3 年前发表的论文, 分别根据文献计量分数排序结果, 为论文定级。定级方法参见本体系 3.1.4 (6)。

## 3.3 综合评估

### (1) 归一化处理

由于同行评议和文献计量指标分数的量纲不同, 需要经归一化处理, 把同行评议指标得分结果  $\bar{U}_j$ , 映射为 (0,1) 内数字, 归一化方法与 3.2.3 (1) 相同。

### (2) 综合得分计算

### 1) 对 1 年评估数据的合成计算

综合得分 =  $0.85 \times$  “同行评议指标得分归一化值” +  $0.15 \times$  “文献计量指标得分归一化值”

### 2) 对 3 年评估数据的合成计算

综合得分 =  $0.7 \times$  “同行评议指标得分归一化值” +  $0.3 \times$  “文献计量指标得分归一化值”

### (3) 分数排序

利用评分软件,分别对论文的 1 年评估数据和 3 年评估数据的综合得分数据,在相同学科内进行排序。

### (4) 论文定级

根据综合得分排序结果,分析得分数值分布规律,为论文定级。定级方法参见本体系 3.1.4 (6) 部分。

## 3.4 评估结果发布

根据论文同行评议和文献计量评估的基本数据,对论文、期刊、机构和作者等评估对象的学术影响力开展相关分析研究并发布成果。评估成果主要分为两类:

- **即年评估:** 对前一年发表的论文的同行评议、文献计量和综合评估结果。
- **3 年评估:** 对 3 年前发表的论文的同行评议、文献计量和综合评估结果。

### **\*指标体系的扩展应用**

本指标体系的实施具备良好的可扩展性，可根据不同需要，对本指标体系所设置的指标进行多种组合对论文进行动态评估。比如，对尚未正式发表的学术论文，可选择学术创新程度、论证完备程度、难易程度、社会价值和课题立项五项指标进行即时的同行评议；对已正式发表论文，若需要即时评估结果，也可在五项指标基

## **第 4 部分：指标体系相关研究报告**

## 4.1

# 人文社会科学论文质量评估指标遴选 文献调研报告

中国人民大学人文社会科学学术成果评价研究中心

2010年7月15日

## 4.1.1 文献调研概述

### 4.1.1.1 研究背景与目标

本报告是“人文社会科学论文质量评估指标体系”设计的系列分报告之一。研究者、科研管理者和学术评价者普遍认同：评价人文社会科学成果是对“评价的再评价”，无论是从价值角度还是方法角度，必须坚持直接指标与间接指标、主观评价与客观评价、学术价值与社会效益有机结合等原则，体现学科共同体的“共识性价值判断”。但是，希望按这些原则建立一个面面俱到万全的评价标准是不可能的，也是不现实的。

构建一个相对科学可行的“人文社会科学论文质量评估指标体系”，首要的问题是，如何充分吸收已有研究成果的精华，将这些抽象原则、共识性价值具体化为较科学的、易操作的评估指标。本报告通过汇总吸收学界、科研管理部门探索的既有成果，特别是总结人大复印报刊资料几十年坚持的选稿标准，归纳梳理出受到广泛认同和普遍共识的人文社科学术论文质量评估的基本指标，并纳入“人文社会科学论文质量评估指标体系”。

### 4.1.1.2 研究内容

本报告通过“文献调研”方法，着重对如下三个方面的一系列问题进行系统全面的文献调研和归纳总结：

#### (1) 人文社会科学论文分类评估方案：

- 已有研究成果中，对人文社会科学论文分类评估的方案包括哪些？

- 这些分类评估方案具有哪些优势和劣势？
- 本研究所构建的“人文社会科学论文质量评估指标体系”应采用怎样的分类评估方案？

## **(2) 人文社会科学单篇论文同行评议基本指标的遴选和设置：**

- 已有研究成果中，受到理论研究者、科研管理者和学术评价者较为普遍认同的评估标准包括哪些？本研究应遵循哪些评估标准设计指标体系？
- 研究者、科研管理者和学术评价者已经提出了哪些人文社会科学论文的同评议指标？这些指标的内涵和外延是怎么界定的？在评价人文社会科学论文质量时有哪些优势和不足？本研究应对这些指标进行哪些改造或创新并纳入本研究所设计的“人文社会科学论文质量评估指标体系”中？
- 现有同行评议指标的逻辑关系如何？是否需要考虑指标的适用性？

## **(3) 人文社会科学单篇论文文献计量评估指标的遴选和设置：**

- 国内外现有评价体系中，正在使用哪些文献计量指标可评估人文社会科学研究成果的质量？
- 这些文献计量指标（如期刊评估指标）是否可直接应用于单篇人文社会科学论文质量的评估，是否需要适当改造或创新设计？
- 在运用文献计量指标评估人文社会科学论文质量时，需要考虑哪些适用条件？

## 4.1.2 论文分类评估方案的调研

研究者、科研管理者和学术评价者普遍主张对人文社会科学学术成果进行“分类评估、同类比较”。大多数人认为，分学科评估是其基本的原则。关于人文社会科学研究成果分类评估的方式，主要有以下三种看法：一是设置相同的指标体系，评估所有成果；二是针对不同的成果，设置不同的指标体系；三是设置通用的评估指标，根据成果的不同类型设置不同的指标权重。关于人文社会科学成果分类评估的探讨，主要包括以下四类理论：

### 4.1.2.1 借鉴自然科学领域，按研究性质（或阶段）分类

1978年，联合国教科文组织在第二十届会议上通过了“关于科技统计国际标准化的建议案”，提出了基础研究、应用研究和试验发展三类科研成果的分类法。我国部分人文社会科学成果的评估借鉴了这一分类法，如北京市哲学社会科学成果评估，将成果分为理论性成果、应用性成果和其他研究成果，分别设置不同的指标体系和权重；《浙江省哲学社会科学优秀成果评奖办法实施细则》中将学术成果分为基础理论类研究成果（包括专著、论文等）、应用类研究成果（包括调研报告、对策建议）、学术资料整理类成果、译著类研究成果四类；《北京大学人文社会科学研究优秀成果评奖办法(试行)》、《黑龙江省高校人文社会科学研究优秀成果奖励办法实施细则》等也采用此类分法。

一些学者也赞成以基础研究和应用研究为主的分类评估方案。比如么大中<sup>28</sup>将科研成果分为理论性和应用性两类，张书伟<sup>29</sup>认为，除此之外，还应有增

<sup>28</sup> 么大中. 关于社会科学优秀成果评选标准问题[J]. 科研管理, 1986(1): 63-66.

<sup>29</sup> 张书伟. 社科学术成果评定的原则及其指标体系[J]. 社科百家, 1991(1).

加“科普资料类”成果。唐德章也认为<sup>30</sup>，可分为理论性成果（抽象性、价值显现时间长）、应用性成果（直接用于实践，时间短）与工具书、资料汇编、译著等三类。这些学者都认为理论性成果应着重评估成果的学术水平和学术影响，应用性成果则着重评估其应用价值和社会影响。因此，基础研究成果的学术价值指标的权重应较大，应用研究成果的社会价值指标应占有较大的权重。丁军强也同意这一观点<sup>31</sup>，并补充认为资料编译类成果真理性指标的权重应最大。

就人文社会科学学术论文的评估而言，基础理论性论文一般是指研究周期长、不直接应用与实践的论文，评估的重点是理论创新；应用实证研究类论文旨在直接用于实践，评估的重点是社会价值。然而，本研究还认为，人文社会科学的许多研究，大多没有明确的推广应用环节，难以在两类之间做出严格区分，因此该分类较难适合人文社会科学论文的评估。此外，理论研究和应用研究成果的评估，应同时体现论文的学术价值和社会价值，差异不在于指标的设置，而在于指标权重的分配。

#### 4.1.2.2 按成果的智力加工深度分类

卜卫，卢渝等人<sup>32</sup>，主张将社科成果分为研究类、普及类和资料类（卢渝认为还有翻译类），各类成果的重要性依智力加工深度递增，分别设置不同的指标体系；杨育华<sup>33</sup>主张按照成果属性将社科成果分为数据资料类、实践类、理论性成果和综合性成果四类，每类成果给予不同的评分范围，分数随智力加工深度递

<sup>30</sup> 唐德章，夏元林. 社会科学研究成果定量评估方法初探[J]. 社会科学研究, 1989 (1): 78-87.

<sup>31</sup> 丁军强，吴桂鸿. 试论社会科学研究成果的评估标准[J]. 科技管理研究, 2007 (6): 87, 91-92.

<sup>32</sup> 卢渝. 社会科学研究成果分类标准及其量化途径[J]. 社会科学, 1992 (4): 73-76.

<sup>33</sup> 杨育华. 试论社会科学研究成果的评估[J]. 宁夏社会科学, 2002 (9): 125-128.

增。罗平安主张<sup>34</sup>，评估过程将成果、论文水平高低分为四大类：最高成果（概念类）、重要成果（模型类）、次要成果（实验与理论计算数据符合类）、一般成果（积累基本数据类）。

若按智力加工深度对学术论文进行分类，综述文章是资料类成果的典型代表。综述文章不像其它研究论文那样以提供新认识为主，而是重在进行新的汇总、概括，为进一步的研究提供资料<sup>35</sup>，和其他论文的差异十分明显。因此理论上讲，综述文章的评估标准应与其他论文存在一定差异。

#### 4.1.2.3 按人文社会科学学科分类

人们普遍认为，人文社会科学学术成果评估应分学科进行同类比较，但尚未有人提出针对各学科特点设计不同的指标体系。国内外著名的评估体系，也都把“按学科分类评估”作为基本的实施原则。蔡曙山的研究<sup>36</sup>将科研成果分为自然科学、工程技术、艺术人文、社会科学 4 个部类共 11 门，建议对艺术人文学科成果和社会科学成果进行分类评估。部分学者认为，人文学科在研究范式上与社会科学存在一定差异，这一差异甚至超过社会科学和自然科学的差异，因此应分为人文学科和社会科学两类进行分别评估<sup>37</sup>。

本研究也认为，应遵循学术共同体的“共识性判断”，对人文社会科学论文的评估坚持“按学科分别评估”的基本原则。但是各学科论文在评估标准、指标、权重等方面的差异性，需要进一步探讨。

<sup>34</sup> 罗平安，成果、论文水平评估标准，非出版物。

<sup>35</sup> 卜卫，周海宏，刘晓红．社会科学成果价值评估[M]．社会科学文献出版社，1999. 5.

<sup>36</sup> 蔡曙山．代表性学术成果作为艺术人文与社会科学评估指标的意义——兼论 CSSCI 的科学评估功能及其发展方向[J]．江西社会科学，2004（6）：208-220.

<sup>37</sup> 庞学铨．论人文科学的价值与功能[J]．中共浙江省委党校学报，2009（2）：11-17.

#### 4.1.2.4 按成果的形式分类

按成果形式分类评估的典型代表是《中国高校人文社会科学研究优秀成果奖励暂行办法》，该办法规定参评成果包括专著、论文、研究咨询报告、译著、考古发掘报告、工具书、古籍整理、软件、音像制品等。当代的部分学者也认为，人文社会科学成果的评估，应首先按照成果形式进行分类。有的学者将其分为学术专著和学术论文，并把教材工具书等纳入著作类，把研究报告等纳入论文类<sup>38</sup>，有的则认为除此之外还包括调研报告<sup>39</sup>，还有人在上述三类之外增加了工具类、译著类、资料类等其他成果<sup>40</sup>。然而，民国时期的学术评估，只把学术性最强的著作和论文作为学术成果，译作、教材、工具书等皆不列入学术成果之列<sup>41</sup>。学术论文是否该单独作为一类，是学者们争论的重要问题。一部分学者认为，学术论文的评估指标在权重上应有别于其他类型的成果，比如学术论文在创新性上要求更高，但与其他成果，尤其是专著相比，体系性和完整性要求低一些<sup>42</sup>。课题来源指标对于学术论文来讲，其重要性比不上研究报告和学术著作；发表层次是学术论文的硬指标，它和“转载与引用”指标构成学术论文的一个重要因素；论文类成果的成熟程度指标应低于著作类成果<sup>43</sup>。另一部分学者认为<sup>44</sup>，学术论文应与学术专著同类，执行相同的指标体系，因为学术论文和学术专著一样，具有

<sup>38</sup> 沙似鹏，郑礼，郭才伯，张毅. 人文、社会科学研究成果评估指标体系初探[J]. 上海高教研究，1994（1）：74-78.

<sup>39</sup> 刘大椿等. 人文社会科学研究成果评估体系研究[M]. 经济科学出版社，2009. 9.

<sup>40</sup> 叶蓬. 人文社会科学研究成果评估指标体系分析[J]. 探求，2001（1）：60-67.

邓毅. 社科成果评估指标体系探析[J]. 现代教育论丛，2004（6）：21-26.

<sup>41</sup> 刘明. 学术评估制度批判[M]. 长江文艺出版社，2006. 1.

<sup>42</sup> 叶蓬. 人文社会科学研究成果评估指标体系分析[J]. 探求，2001（1）：60-67.

邓毅. 社科成果评估指标体系探析[J]. 现代教育论丛，2004（6）：21-26.

<sup>43</sup> 叶蓬. 人文社会科学研究成果评估指标体系分析[J]. 探求，2001（1）：60-67.

邓毅. 社科成果评估指标体系探析[J]. 现代教育论丛，2004（6）：21-26.

<sup>44</sup> 杨霞，熊世春. 社会科学研究成果综合评估指标体系研究[J]. 山西师大学报(社会科学版)，2008（5）51-54.

卜卫，周海宏，刘晓红. 社会科学成果价值评估[M]. 社会科学文献出版社，1999. 5.

较强的创新性和学术性，但实践性和应用性较其他成果较差。

#### 4.1.2.5 结论分析

对上述典型的分类评估方案的优劣势分析详见表 10。

表 10 现有分类评估方案分析

分类标准	分类结果	差异性	优劣分析
按研究性质	理论性成果（也称基础研究成果）	注重评估成果的学术价值，创新性、科学性指标的权重较大	<ul style="list-style-type: none"> <li>➢ 有一定区分度和合理性，但有些论文的理论性和应用性较为模糊，比如综述文章。</li> <li>指标应同时体现学术价值和社会价值，差异在于指标的权重。</li> </ul>
	应用性成果（也称实践成果）	注重应用效益和社会影响，社会价值指标的权重较大	
	其它成果：如翻译类、整理类、普及资料类、工具类等	学术论文中，只有综述文章可划归此类	
按智力加工深度	研究类（重要）	提供新知识	<ul style="list-style-type: none"> <li>➢ 就论文而言，可分为一般研究类和综述文章两种。</li> <li>➢ 区分度较高，具有一定合理性。</li> <li>➢ 能够涵盖所有论文。</li> </ul>
	普及类（中等）	不提供新知识，对原有知识提取、概括、加工和转述，学术论文中，只有综述文章可划归此类	
	资料类（次要）	为研究提供基本工具、资料、信息和文献，不提供新知识，不适用于学术论文	
	其它类：如翻译类、	不适用于学术论文	

分类标准	分类结果	差异性	优劣分析
	综合类		
按学科大类	人文类	以人的精神世界为研究对象的论文	<ul style="list-style-type: none"> <li>➢ 涵盖度较好，但人文社会科学学科划分概念较为模糊，较难区分。</li> <li>➢ 按学科分类，适用于评估实施中进行同类比较。</li> </ul>
	社科类	以社会现象为研究对象的论文	
按成果类型	专著	注重评估系统性、完整性、成熟度	<ul style="list-style-type: none"> <li>➢ 论文与其他形式的成果相比，特征突出，有必要单独设置指标体系进行评估。</li> <li>➢ 本研究只针对学术论文，因此应更为关注创新性和发表层次。</li> </ul>
	论文	注重创新性和发表层次，系统性、完整性、成熟度稍差；课题来源的重要性不及研究报告	
	研究报告	注重课题来源	
	其它：教材工具、译著、古籍整理、音像成果等	与本研究无关	

综上所述，本研究认为：1) 按学科进行分类评估、同类比较，得到学术共同体和实践者的广泛认同；2) 对论文的分类评估，差异主要体现在权重设置，

而非指标设置上；3) 综述文章与研究论文存在明显差异，需要给予重点关注；4) 学术论文在形式上具有“短小精干”的特征，在构成要素上也自成体系，因而有必要作为单独一类设置评估指标体系。

### 4.1.3 同行评议指标的调研

学界普遍认为人文社会科学成果评估应该更加关注内容质量(而非形式和数量)。本部分从人文社会科学论文质量评估的需要出发，对可用于人文社会科学论文质量同行评议的评估标准进行汇总和探讨，对同行评议指标的设置进行分析和选择，对指标的科学性、适用性和逻辑关系进行论证。其中，同行评议指标将分为内在指标和外在线标<sup>45</sup>进行阐述。

#### 4.1.3.1 论文质量的评估标准

评价标准是评价主体评估时所要遵守的基本理念和操作标准，贯穿于评价指标体系构建的全过程。因此，我们对当前研究者、科研管理者和学术评价者提出的评估标准进行了汇总和分析（详见表 11）。

表 11 关于成果评估标准的文献分析

评估标准	含义	认同程度	涉及的指标	本研究的观点
创新标准	考察成果在内容上的新意和对学术发展的贡献	一致认同	创新程度、新颖性等	应作为核心标准

<sup>45</sup> 丁军强等. 社会科学研究成果评估指标体系构建之管见[J]. 评估与预测, 2006(4): 113-115.

评估标准	含义	认同程度	涉及的指标	本研究的观点
科学标准	包含逻辑标准；考察研究过程是否正确、可靠	普遍认同	科学性、完备程度、成熟程度、逻辑性、可靠性、规范性、先进性等	应作为核心标准
价值标准	包括学术价值和社会价值，考察成果对于学术共同体和社会的有用性	普遍认同	学术价值、社会价值、社会影响、社会效益、政府采纳、媒体报道等	与创新标准和科学标准有重合，评估角度不同。
效益标准	考察成果对促进经济发展的作用	大部分认同	社会效益、效益价值、经济效益等	与价值标准中的社会价值存在重合，侧重反映成果对经济发展的价值。
学术标准	考察成果的“学术含量” <sup>46</sup> ；与非学术标准相对	部分认同	学术性、学术规范性等	“学术性”是隐含的前提条件。“学术标准”与“非学术标准”是当前评估中待深入研究的难点之一 <sup>47</sup> 。
政治标准 <sup>48</sup>	考察成果的与执政党思想的一致性	部分认同	政治性、政府采纳等	作为本体系评估的前提条件，不符合该标准的论文不纳入评估范围。

<sup>46</sup> 贾岩. 编辑学视角中的社科论文评估[J]. 湖南大众传媒职业技术学院学报, 2006 (1): 82-83.

<sup>47</sup> 刘大椿. 中国人文社会科学评估问题之审视[J]. 重庆大学学报(社会科学版), 2009 (1): 54-59.

<sup>48</sup> 贺忠德. 社会科学研究成果鉴定标准和评估方法初探[J]. 社科论坛, 1998 (1-2): 66-71.

评估标准	含义	认同程度	涉及的指标	本研究的观点
审美标准 <sup>49</sup>	考察成果的可读性， 比如阐述清晰程度、 语言的优美程度等	少部 分认 同	可读性、阐述清晰程 度等	与学术论文质量相关度不 大，本研究不考虑纳入。

(注：创新标准和科学标准两者是真理标准的核心体现)

总体上，“真理标准”和“价值标准”是学界普遍认同的两大评估标准。学界普遍认为应对不同评估标准进行综合考虑。鉴于表 2 中所列的上述评估标准，存在一定的重叠，因此本研究在设计指标体系时，综合考虑被学界广泛认同的标准，确定将评估标准确定为“真理标准”为主，“社会价值标准”为补充。

#### 4.1.3.2 同行评议的内在指标

内在指标是指论文内容本身所反映出的、与质量直接相关的指标。

##### (1) 创新性（也称创新程度、创新力度、新颖性等，涉及学术价值）

在内涵上，（学术）创新性包含三层含义：1) 与以往不同；2) 提供有价值的知识；3) 这种“不同”限于“某一学术课题”领域之内<sup>50</sup>。在外延上，创新性还可分解为新领域（新选题）、新应用、新资料、新布局、新概念、新方法、新观点、新视角、新论证、新的总结概括等。现有研究普遍认为，创新性作为评估指标，在内涵上不再细分二级指标，在外延上各种不同的创新应为并列关系，具备其一即可认为具有创新性。比如国家社科基金项目成果评估指标体系、辽宁

<sup>49</sup> 胡红生. 人文社会科学成果的评估标准[J]. 武汉理工大学学报(社会科学版), 2001(8): 316-319.

<sup>50</sup> 周露阳. 学术论文的创新途径: 基于创新因素的分析[J]. 科技管理研究, 2006(8): 195-197.

省财政科研基金项目成果评估指标体系、北京市哲学社科科学成果评估指标体系等都认为，成果具备任何一项创新都被视为具有创新性，评估是针对创新的程度进行的。然而，难以找到较为清晰的方式描述创新程度。周露阳认为<sup>51</sup>创新有“大小”之分，“小创新”是指局部与渐进的创新；“大创新”是整体与彻底的创新。

创新被誉为繁荣人文社会科学的灵魂<sup>52</sup>，该指标几乎得到国内外学者的一致认可，适用于评估所有学科和类型的论文。对综述文章而言，创新性着重于新的总结概括，而不是新的理论或方法，所以通常只能实现“小创新”。本研究主张将创新性作为指标体系的核心指标，并且不再细分二级指标。

## (2) 科学性 (也称真理性、完备性、成熟度等, 涉及学术规范性)

已有研究对科学性指标内涵和外延的理解不甚一致。主要有三种观点：1) 科学性评估是指对成果认识的正确性和真理性进行的判断<sup>53</sup>，可细分为可靠性、逻辑性和学术规范性三个二级指标<sup>54</sup>；2) 学术规范性应单独作为指标，不属于科学性范畴<sup>55</sup>；3) 科学性指标可用来衡量成果的研究过程是否充分合理和严密可靠，可细分为理论前提科学性、资料全面充分性、研究方法有效性、概念使用准确性、逻辑推理严密性、学术引证规范性等五个二级指标<sup>56</sup>。此外，美国的评估则认为延续性也是“真理标准”的重要内容，因为连续的研究更有利于

<sup>51</sup> 周露阳. 学术论文的创新途径：基于创新因素的分析[J]. 科技管理研究, 2006(8): 195-197.

<sup>52</sup> 吴树青. 创新是繁荣哲学社会科学的灵魂[N]. 光明日报, 2006-01-16.

<sup>53</sup> 吴桂鸿. 社会科学研究成果评估指标体系研究[D]. 湖南大学, 2006.

<sup>54</sup> 注释：如《国家社科基金项目成果评估体系》、《河南省社科规划项目成果鉴定评估体系》等规定的“完备程度”，以及叶蓬的研究所界定的“成熟程度”。

<sup>55</sup> 注释：如社科院卜卫的研究所界定的“科学性”，《北京市哲学社会科学项目成果评估指标体系》等。《北京市哲学社会科学成果评估体系》还认为科学性还应包含系统性的内涵，但本文认为系统性对于学术论文来讲可以弱化，因此未单列。

<sup>56</sup> 卜卫, 周海宏, 刘晓红. 社会科学成果价值评估[M]. 社会科学文献出版社, 1999. 5.

全面占有信息、避免重复劳动，易得出更为准确的结果<sup>57</sup>。各类研究体系普遍认为，评估是针对研究过程的科学程度进行的，只有上述所有内容满足了科学性要求，整个成果才具备了科学性；局部的欠缺和不足，将减少成果的科学性。

科学性指标是学术论文真理性标准的重要体现，国内外学者普遍认为，该指标对于评估学术论文的质量非常重要，适用于评估各学科和类型的论文。本研究主张，将科学性指标纳入指标体系。虽然当前对科学性指标的认识不甚一致，但我们认为凡是对成果论证的正确性有影响的，都应包含在科学性指标之内。在构建指标体系时，如何界定科学性指标的名称和内涵，细分二级指标，需要仔细斟酌。综述文章科学性的评估，尤其应注重资料的充分性和真实性，但科学性指标的其他内容同样也适用，差异不大。

### **(3) 学术价值（涉及创新性、科学性）**

现有研究对这一指标的界定较为模糊和混乱。笼统地讲，学术价值是指论文对学科发展的贡献程度<sup>58</sup>。对于学术价值指标的分解，有许多不同看法：有的主张细分为创新性、成果成熟度、学科建设贡献、对学术界和社会的影响、论文发表刊物等级等<sup>59</sup>；有的主张细分为专门性、系统性和创新性<sup>60</sup>，还有的主张细分为新理论、新观点、新方法、新论证等<sup>61</sup>。在指标体系设置上，有的学者主张将学术价值与创新性、科学性指标并列设置，比如北京市哲学社会科学成果评估指标体系，也有的主张将创新性和科学性作为学术价值的二级指标或组成内容，还

<sup>57</sup> 国家社会科学院外事局. 美国人文社会科学现状与发展 (M). 北京: 社会科学文献出版社, 2001: 370-391.

<sup>58</sup> 注释: 《北京市哲学社会科学成果评估指标体系》.

<sup>59</sup> 唐德章, 夏元林. 社会科学研究成果定量评估方法初探[J]. 社会科学研究, 1989 (1): 78-87.

<sup>60</sup> 郭开选. 文科学术论文学术价值的估量[J]. 渭南师范学院学报, 2002 (7): 90-92.

<sup>61</sup> 卜卫, 周海宏, 刘晓红. 社会科学成果价值评估[M]. 社会科学文献出版社, 1999. 5.

有学者认为，学术价值就等于创新性，比如社科院卜卫的研究。

有相当一部分学者认为，评估指标体系中应包含“学术价值”指标。本研究也认为，具备学术价值是一篇合格学术论文的基本要求；即使是以实践应用为目的的论文在学术价值上可以稍微弱化，但仍需要充分体现学术价值。学术价值是指论文对学术共同体的有用性，集中体现在提供新的观点、方法等，但论证过程和方法对其他学者也会产生启发性、参考性等作用，因此科学性指标也是学术价值的重要体现。

由于“学术价值”这一概念，在作为评估指标时，内涵过于模糊，外延难以清晰界定；且创新性、科学性等其它指标能够较为有效地反映论文的学术价值，因此本研究主张，不将其作为独立评估指标列入指标体系，而是通过其他指标反映学术价值。

#### **(4) 社会价值（也称社会影响、社会反响、社会效益、经济效益等）**

“社会价值”指标，用以衡量学术成果对于推动现实问题解决的积极作用<sup>62</sup>。社会价值体现在成果对经济发展、社会发展、思想道德进步、政治观点、生活方式等诸多方面带来的积极影响。还有学者认为，被引用转载情况、被采纳情况、获奖情况、媒体报道等，是反映论文社会价值的重要因素（引用），但是较难对社会价值指标进行二级分解，因为从外延上进行列举，是难以穷尽的。目前主要的观点包括：1) 把“成果价值”设置为一级指标（与创新指标并列），把学术价值、社会价值等作为“成果价值”的二级指标，并规定具有二级指标之一即可认

<sup>62</sup> 注释：《国家哲学社会科学项目成果评估指标体系》。

为具有价值性<sup>63</sup>，用来评估成果价值的含量大小。2) 应用成果评估指标体系中，把社会价值、经济效益和政府采纳情况作为并列指标同时进行评估<sup>64</sup>。3) 将“社会价值”指标细分为经济效益（考察投入产出因素）、社会效益、对决策和管理的作用等二级指标。4) 将“社会价值”指标细分为政治价值、市场价值和伦理价值等二级指标<sup>65</sup>；5) 应对理论成果评估其“社会价值”，对应用成果评估其“社会效益”。

国内外研究普遍认为，人文社会科学研究成果的价值，不仅仅在于对社会现象的深刻认识，更在于它解决社会问题、推动社会向前发展的能力。然而从以上分析可以看出，当前对社会价值内涵和外延的理解十分混乱。本研究认为。“社会价值”是指，学术成果超出学术范畴、对于社会其它方面的有用性。所有的论文都具有一定的社会价值。

由于社会价值是学术论文质量的重要体现之一，并且社会价值难以通过学术创新程度、科学性等指标得以反映，而是需要从社会的多元化实际需要出发，对论文的社会价值进行单独判断，因此，本研究主张将社会价值指标纳入指标体系。然而，由于“社会价值”指标的特殊性，在评估时具有诸多理论难点，需要在指标体系构建过程中着重关注：

第一，“社会价值”指标，评估的是社会价值的含量还是实现程度？可以准确预测的吗？若用于同行评议及时评估，只能对成果的社会价值含量进行预测。成果的社会价值具有“隐性”和“显性”之分：有些应用实证研究成果（比如政

<sup>63</sup> 注释：如国家和北京市《哲学社会科学项目成果评估指标体系》等。

<sup>64</sup> 邓毅，社科成果评估指标体系探析[J]，现代教育论丛，2004（6）：21-26。

<sup>65</sup> 谭贤楚，刘伦文，叶明，人文社会科学成果评估：必要性、理论及方法——基于社会学的分析与思考[J]，湖北民族学院学报（哲学社会科学版），2006（2）：155-158。

策研究)与现实联系紧密,较容易预测其社会价值;有些基础研究成果,与现实存在一定疏离(引用),加之其社会价值的实现过程是隐性的,无疑给社会价值的判断带来了困难。但是社会价值的是否明显,只能说明评估的有效性,并不能说明社会价值的含量。因而确保社会价值判断的有效性十分重要,同行评议只能对成果“可预测”的社会价值进行判断。换句话说,同行评议及时评估对社会价值的判断是有限的,因此该指标的权重不应过高。

第二,如何界定社会价值指标的名称、内涵和外延:从价值角度评估还是从效益角度评估?从哲学角度进行价值评估还是从经济学角度进行价值评估?是否需要对该指标进行细化?从前文所述的文献研究可以看出,社会价值的内涵极为丰富,对于该指标的分解也未形成有效的观点。因此,本文初步认为:采取同行评议对论文的社会价值进行及时评估时,只能对该指标的含义进行“模糊处理”,不再细分二级指标,而以定性的指标描述引导评估者利用该指标进行评估。

第三,学术成果社会价值的大小与社会的多元化需求密切相关,应由学术共同体还是社会群体来评估论文的社会价值?对不同主体的社会价值具有可比性吗?很明显,由利益相关的社会群体来判断论文的社会价值,更为科学,然而可操作性却很差,因为利益相关群体的组成过于复杂。因此,本研究主张仍由学科领域专家作为评估主体,对论文的社会价值进行预测。所预测的内容是抽象的“价值”,因而可认为具有可比性。

### (5) 难度 (也称难易程度、难度系数等, 涉及重要性)

通常来讲,“难度”指标用来衡量成果的“平均劳动量”<sup>66</sup>;在其他指标评估获得同等结果的前提下,难度越大表示质量越高。有些指标体系还把“难度和重要性”设置为同一指标。对于难度指标的理解主要包括以下三种观点:1)大多数研究者认为,“难度”指标可细分为“研究难度(复杂性)”和“资料搜集处理难度”两个二级指标<sup>67</sup>;2)个别学者认为<sup>68</sup>可细分为“研究规模”和“复杂程度”两个二级指标;3)还有人认为<sup>69</sup>应细化为“新颖程度”、“复杂程度”、“艰巨程度”和“获取难度”四个指标。此外,有学者认为理论研究成果和应用研究成果在难度的考察上,指标设置应有所差异。比如唐德章认为<sup>70</sup>,难度主要是针对理论性成果而言的,可细分为课题来源、涉及学科领域(含跨度和基础)、需求程度、风险程度四个二级指标;贺忠德则认为<sup>71</sup>,理论成果的“重要性和难度”可以从项目类别、研究难度、实践程度、逻辑水平四个方面评估,应用类研究成果则需要考察项目类别、需求程度和研究难度三个方面。关于难度指标与其他指标关系的问题,主要有两种观点:1)“难度”指标与创新性、科学性等指标属于并列关系,结果可进行相加。2)“难度”应作为系数,与其他指标结果相乘<sup>72</sup>。

“难度”指标虽并未纳入到所有的指标体系,但也有相当一部分学者认同该指标对于评估成果质量的重要意义。本研究认为,难度指标是反映论文质量的重

<sup>66</sup> 卜卫,周海宏,刘晓红. 社会科学成果价值评估[M]. 社会科学文献出版社,1999.5.

<sup>67</sup> 注释:如国家和北京市、河南省、辽宁省等《哲学社会科学评估指标体系》。

杨育华. 试论社会科学研究成果的评估[J]. 宁夏社会科学,2002(9):125-128.

<sup>68</sup> 卜卫,周海宏,刘晓红. 社会科学成果价值评估[M]. 社会科学文献出版社,1999.5.

<sup>69</sup> 卜卫,周海宏,刘晓红. 社会科学成果价值评估[M]. 社会科学文献出版社,1999.5.

喻承久,张勇. 社会科学成果评估指标体系分析[J]. 空军雷达学院学报,2005(9):73-75.

<sup>70</sup> 唐德章,夏元林. 社会科学研究成果定量评估方法初探[J]. 社会科学研究,1989(1):78-87.

<sup>71</sup> 贺忠德. 社会科学研究成果鉴定标准和评估方法初探[J]. 社科论坛,1998(1-2):66-71.

<sup>72</sup> 喻承久,张勇. 社会科学成果评估指标体系分析[J]. 空军雷达学院学报,2005(9):73-75

要方面，其他指标无法替代，因此应纳入本研究的指标体系。不同学科和类型的研究成果，在难度上可能存在差异，因而难度应在相同或类似的学科成果内部通过比较确定。与其他指标相比，难度指标的内涵更为抽象，因而如何界定该指标的内涵和外延，还需重点关注。

#### 4.1.3.3 同行评议的外在指标

外在指标不是论文内容本身所反映的，但能间接反映论文质量的指标。

关于外在指标，学界的观点是针锋相对的：**1) 完全否定外在指标对于评估学术成果质量的作用。**具有代表性的就是社科院卜卫、陈力丹等，其中卜卫的社会科学成果评估体系被称为“五不评估体系”<sup>73</sup>，完全不考虑任何外在指标，在同类研究中独树一帜。在成果评估实践中，国家和各省的社科成果评估体系，因主要采用同行评议，也不考虑外在指标。**2) 把外在指标作为评估成果质量的重要指标。**持这一观点的以科研管理者和部分学者<sup>74</sup>为主要代表，他们认为外在指标是评估成果质量的重要依据。因此，应将课题来源、发表层次、成果字数、被转载应用情况、获奖情况、引用情况等，作为评估成果质量的重要指标，有学者指出<sup>75</sup>，对于单篇论文而言，首要评估层次是发表级别。

随着研究的深入，当前大多数学者更倾向于内在指标和外在指标两者相结合的观点。比如杨霞、熊世春认为<sup>76</sup>，项目级别、引用情况等可作为评估时的“附加

<sup>73</sup> 注释：五不体系是指：①不以课题来源评成果；②不考虑媒体的报道和评估；③不承认所谓“核心期刊”；④不以评奖论英雄；⑤学术同行的引用率暂时无法作为衡量标准。

<sup>74</sup> 叶蓬．人文社会科学研究成果评估指标体系分析[J]．探求，2001（1）：60-67．

<sup>74</sup> 沙似鹏，郑礼，郭才伯，张毅．人文、社会科学研究成果评估指标体系初探[J]．上海高教研究，1994（1）：74-78．

杨家栋，秦兴方．社会科学研究成果的评估及其指标体系[J]．齐鲁学刊，2001（2）：122-128．

<sup>75</sup> 雷二庆，单篇科技期刊论文评估方法初探[C]，第十届全国科技评价学术研讨会精选论文集，2010．

<sup>76</sup> 杨霞，熊世春．社会科学研究成果综合评估指标体系研究[J]．山西师大学报（社会科学版），2008（5）

分”。丁军强、吴桂鸿等认为<sup>77</sup>，人文社会科学成果的评估应将外在指标与内在指标相结合，外在指标和内在指标的权重，在评估成果质量时应“三七开”。

### **(1) 课题来源（也称课题立项、项目来源、选题来源、课题资助等）**

指资助论文研究的课题或项目。大部分研究认为可将课题分为海外资助课题、国家级课题、其他国家级课题、省级课题、省级以下课题等不同级别；也有学者<sup>78</sup>将课题按照经费规模和数量进行划分，比如10万以上，5-10万元以上，2-5万元等。不管采用哪种分级方式，都存在不足：前者无法用来划分横向课题和商业课题，后者的分级中，经费不能代表课题质量，因而缺乏充足依据。“课题来源”指标或用来表示成果的难度和重要性<sup>79</sup>，或用来表示学术创新性和学术价值<sup>80</sup>。目前国内省级以上的人文社会科学课题资助的项目，评审标准和流程类似，但其他类型的课题资助评审各自为政。

赞成和反对将课题来源列入指标体系的争论较为激烈。反对者主要是认为课题来源与成果质量不能完全形成正比；赞成者则认为课题级别越高或经费越多，产生较高质量论文的可能性越大，能够有效反映成果质量。关于“课题来源”指标的权重，有的学者认为<sup>81</sup>课题来源指标的权重分配，对于研究报告和学术专著来说，应高于学术论文。

51-54.

<sup>77</sup> 丁军强等. 社会科学研究成果评估指标体系构建之管见[J]. 评估与预测, 2006(4): 113-115.

吴桂鸿. 社会科学研究成果评估指标体系研究[D]. 湖南大学, 2006.

<sup>78</sup> 沙似鹏, 郑礼, 郭才伯, 张毅. 人文、社会科学研究成果评估指标体系初探[J]. 上海高教研究, 1994(1): 74-78.

<sup>79</sup> 吴桂鸿. 社会科学研究成果评估指标体系研究[D]. 湖南大学, 2006.

唐德章, 夏元林. 社会科学研究成果定量评估方法初探[J]. 社会科学研究, 1989(1): 78-87.

<sup>80</sup> 叶蓬. 人文社会科学研究成果评估指标体系分析[J]. 探求, 2001(1): 60-67.

<sup>81</sup> 叶蓬. 人文社会科学研究成果评估指标体系分析[J]. 探求, 2001(1): 60-67.

“课题来源”指标得到大多数指标体系的赞同。本研究认为，“课题来源”指标在以下两方面能够反映论文质量：第一，大多数情况下，获得课题资助的级别越高、经费越多，获取的研究条件就越充分，越有利于形成高质量的研究成果；第二，大多数情况下，课题立项级别越高，评审越严格，越有利于保障成果的质量。因此拟将这一指标纳入进一步考察的范围内，但难点在于：1) 如何统一不同课题来源的量纲；2) 如何对该指标进行量化评估。

## **(2) 发表层次 (也称期刊级别, 收录情况等)**

“发表层次”通常可理解为期刊的级别。通常认为，期刊级别越高，论文质量越高。目前国内外对期刊进行分级评估体系很多，国际很有影响的比如 SCI、SSCI、AHCI 等来源期刊，国内有影响的比如南京大学“CSSCI 来源刊”、武汉大学“中国学术期刊评估”、中信所《中国科技期刊引证报告》、北京大学《核心期刊要目总览》、社科院《中国人文社会科学核心期刊要览》等，有些评估将期刊更细致地分为 A、B、C 等级别。上述期刊评估体系在期刊选择范围、评估标准、评估机制上各不相同。但总体上期刊的级别划分，存在一定的耦合度。

赞成将“发表层次”作为指标的认为，发表层次越高，对论文质量的控制更严格，因而期刊级别在一定程度上能够反映论文质量。反对者则认为期刊的质量和论文的质量不存在绝对关系，存在“差刊出好文，好刊有差文”的情况。

该指标得到大多数指标体系的赞同。本研究认为，虽然期刊级别不能与论文质量完全成正比，但是大多数情况下，期刊级别越高，评审越严格，发文难度越大，因而能期刊级别能有效反映论文质量。因此拟将这一指标纳入进一步考察的

范围内。但是采用“发表层次”作为评估指标，难点在于：1) 选择哪一种期刊评估体系作为参照标准；2) 如何对该指标进行量化评估。

### (3) 政府采纳情况

指政府对成果的采纳情况：包括采纳部门的级别、成果被采纳的程度、采用后的效益等。采纳的级别可分为国家级、省级、地市级等；采纳的程度可分为全部采纳、部分采纳等；采纳后的效益需要具体计算。政府采纳通常作为反映成果“社会价值”的指标。

这一指标得到部分学者的赞同<sup>82</sup>，也受到了一些科研管理部门的青睐。然而政府采纳的情况只可能出现在以政府为读者对象的科研成果里(如政策研究类论文)；其余的论文，尤其是人文类理论研究论文，针对的读者对象不是政府，很难被管理机构获取，其内容也难以支撑管理机构的需要，因而很难被采纳。所以说是否被政府采纳取决于成果内容、针对群体、转化途径、采纳时间等多种因素，具有很大的偶然性；若作为评估指标，容易造成结果有失公正。因此，本研究不予考虑。

### (4) 被引用情况

“被引用情况”指标依赖引文数据库等统计工具，以及所设置的一系列指标(如被他引数、被引广度、H指数、半衰期等)间接反映论文的质量。通常认为，被引次数越多，级别越高，论文质量越高。该指标需要长期、大量的数据积

<sup>82</sup> 贺忠德. 社会科学研究成果鉴定标准和评估方法初探[J]. 社科论坛, 1998(1-2): 66-11.  
唐德章, 夏元林. 社会科学研究成果定量评估方法初探[J]. 社会科学研究, 1989(1): 78-87.  
沙似鹏, 郑礼, 郭才伯, 张毅. 人文、社会科学研究成果评估指标体系初探[J]. 上海高教研究, 1994(1): 74-78.  
吴桂鸿. 社会科学研究成果评估指标体系研究[D]. 湖南大学, 2006.

累才能得出较为可信的结果，难以用于单篇论文的同行评议及时评估。本研究拟将被引用情况作为引证计量指标，辅助同行评议指标体系的应用（详见本报告文献计量指标部分）。

### **(5) 获奖情况**

指论文获奖的层级。获奖是对论文质量的再次认可和确认，一般被认为是体现论文“社会反响”的指标。获奖级别可分为国际、国家级、省级、校级等，获奖层次越高则表示论文质量越高。

这一指标目前只得到少部分研究者和部分科研管理部门的赞同。与“被转载情况”类似，国内外的学术获奖活动五花八门，评估标准不一。若将获奖情况作为评估指标，容易造成论文评估的不公正。因此，本研究不予考虑该指标。

### **(6) 会议入选**

指依据论文入选会议的情况进行评估。会议可分为国际、国家级、省级等，会议层次越高则表示论文质量越高。会议入选指标被认为是体现论文“社会价值”的指标。只有个别研究者和指标体系赞成把会议入选作为评估人文社会成果指标<sup>83</sup>，本研究认为，国内外各级学术会议收录论文的标准多样，会议入选不能有效反映论文的质量，因此本研究不拟考虑该指标。

### **(7) 成果容量**

指成果的字数。一般来说，成果的字数越多，容量越大，说明成果的质量越

<sup>83</sup> 叶蓬. 人文社会科学研究成果评估指标体系分析[J]. 探求, 2001 (1): 60-67.

贺忠德. 社会科学研究成果鉴定标准和评估方法初探[J]. 社科论坛, 1998 (1-2): 66-71.

高，通常被认为是反映成果“学术价值”和“社会影响”的指标。然而该指标明显是从“数量评估”出发的，不能有效反映论文质量。只有个别研究者和指标体系<sup>84</sup>，赞成将成果容量纳入评估指标体系。随着学术评估“从数量向质量”的转变，本研究认为，不宜将该指标纳入指标体系。

## **(8) 媒体报道**

媒体报道被认为是对论文质量的再次肯定和确认，作为成果的“社会反响”指标，主要受到科研管理部门的青睐。然而学者们普遍反对这一指标，认为媒体报道存在一定的宣传因素，不能有效反映学术成果的质量<sup>85</sup>。因此，本研究不将该指标纳入考虑范围。

### **4.1.3.4 结论分析**

**第一，学界和科研管理部门对评估标准和主要指标已形成较为一致的认识；可用于同行评议的指标，主要集中在内在指标上；适于同行评议的指标，应以内在指标为主，外在指标为辅。得到较多认同的指标包括：创新性、科学性、学术价值、社会价值、难度、课题来源和发表层次。**

**第二，现有研究对上述指标的关系的阐述较为混乱，需要我们在构建指标体系的时候仔细斟酌、加以修正。单纯从概念上来说，层次混乱不合理。比如王贯中<sup>86</sup>的研究，把创新程度分解为可靠性、逻辑性、创新性，明显存在二级指标之和大于一级指标的情况。再比如，有的学者把“创新程度”和“学术价值”并列**

<sup>84</sup> 沙似鹏，郑礼，郭才伯，张毅. 人文、社会科学研究成果评估指标体系初探[J]. 上海高教研究，1994（1）：74-78.

<sup>85</sup> 陈力丹. 谈谈人文社会科学成果的评估标准[J]. 华中科技大学学报（社会科学版），2003（1）：14-17.

<sup>86</sup> 王贯中. 高校科研成果评估指标体系的构建[J]. 科技促进发展，2008（8）：43-46.

设置,有的则认为创新程度即学术价值的核心体现(专指基础研究类)<sup>87</sup>。

**第三,针对不同类型的论文,个别指标的内涵和外延的界定,仍需要精细化。**

比如基础研究成果的社会价值较难预测,应着重考察它的学术价值;应用研究成果的“社会价值”指标的权重应稍高。

综上所述,根据文献调研结果,本体系入选的指标和弃用的指标分别详见表12、13。指标遴选原则包括:1)必要性(普遍认同程度):理论上和实践上得到广泛认同,是遴选指标的首要原则;2)有效性:与论文质量存在正相关,能有效反映论文质量;3)相对独立性:在逻辑上与其他指标不重合,其他指标不可替代;4)可操作性:适于同行评议及时评估。

表 12 根据文献调研入选的同行评议指标

类别	指标	必要性	有效性	相对独立性	可操作性
内在 指标	创新性	广泛认同,适于所有论文	首要指标	不可替代	适于同行评议
	科学性	广泛认同,适于所有论文	关键指标	不可替代	适于同行评议
	社会价值	大部分认同,更适合社会价值较为明显的论文	重要指标	不可替代	适于同行评议
	难度	大部分认同,适于所有论文	重要指标	不可替代,可作为其他内在指标的系数	适于同行评议

<sup>87</sup> 廖萍. 高校社会科学成果评估指标体系研究[J]. 科技管理研究, 2008(5): 169-170.

外在 指标	课题 来源	部分认同，体现了立项 评估者的评估结果	辅助指标	综合体现内在指 标评估结果	可作为参考
	发表 层次	大部分认同，适于所有 论文，体现了期刊编辑 对论文的评估结果	辅助指标	综合体现内在指标 评估结果	可作为参考

表 13 根据文献调研弃用的指标

类别	具体指标	弃用原因
内在 指标	学术价值	<ul style="list-style-type: none"> <li>● 与创新性和科学性有重叠</li> <li>● 不符合指标体系构建标准</li> </ul>
	成熟度、系统性、可靠性、规范性、 先进性、充分性、完备性等	<ul style="list-style-type: none"> <li>● 隶属于科学性，或与科学性重叠</li> </ul>
	唯一性、新颖度	<ul style="list-style-type: none"> <li>● 隶属创新性</li> </ul>
	作用强度、影响力度、效益性	<ul style="list-style-type: none"> <li>● 隶属社会价值或学术价值</li> </ul>
	研究内容意义与前沿性	<ul style="list-style-type: none"> <li>● 与社会价值或学术价值相关</li> <li>● 不能有效反映论文质量</li> </ul>
	可读性	<ul style="list-style-type: none"> <li>● 与科学性有关，但对质量影响小</li> </ul>
	生命强度、影响广度	<ul style="list-style-type: none"> <li>● 不能直接反映论文质量，难以用于同行</li> </ul>
外在	资料性、程序性、功能性、管理性、 理论性、综合性、专门性	<ul style="list-style-type: none"> <li>● 概念过于模糊</li> <li>● 与论文质量不直接相关</li> </ul>
	被转载情况、被引用情况	<ul style="list-style-type: none"> <li>● 不适用于同行及时评估</li> </ul>

类别	具体指标	弃用原因
指标	政府采纳、获奖情况、媒体报道	<ul style="list-style-type: none"> <li>● 不适用于同行及时评估</li> <li>● 具有偶然性，容易导致评估有失公正</li> </ul>
	会议入选、成果容量	<ul style="list-style-type: none"> <li>● 不能有效反映成果质量</li> </ul>

文献调研的同时，研究小组对中国人民大学书报资料中心 50 余年的选文经验进行了实地调研，以期获得实践部门对构建指标体系的一手资料。

表 14 文献调研结果与书报资料中心调研结果对比

文献调研的评估指标	书报资料中心评估标准
创新性	创新，同类比较中是精品，首要指标
科学性	立论、规范
社会价值	体现在同类比较中
难度	体现在同类比较中，在同类研究中有一定深度
课题来源	作为辅助指标
发表层次	作为辅助指标

从表 14 的对比，可以看出，虽然文献调研评估指标和书报资料中心评估标准的概念称谓有所不同，但两者所遵循的评估标准，在理论层面仍存在一定的耦合度，其中创新性、科学性、课题来源和发表层次四个指标的耦合度尤为高。根

据上述文献调研结果,确定了6个指标进行进一步的问卷调查,见4.2部分“问卷二”。

#### **4.1.4 文献计量指标的调研**

文献计量指标是通过引证等量化数据进行统计分析进而评估论文的指标。

##### **4.1.4.1 国内五大文献计量评估体系对比分析**

我国早已大规模引入引证计量方式作为学术评估的重要方式,知名的评估体系包括武汉大学“中国学术期刊评估”,中信所《中国科技期刊引证报告》,北京大学《核心期刊要目总览》,南京大学《中国人文社会科学期刊学术影响力报告》,社科院《中国人文社会科学核心期刊要览》等。对这五大体系所运用的引证计量指标,对比分析如表15所示:

表 15 国内现有文献计量评估指标体系对比

名称	选文范围	数据源获取	评价指标		指标权重分配	创新之处	评价角度
			定量	定性			
武汉大学：中国学术期刊评价	专业类期刊 4665 种和高校学报 1464 种	前四个指标的数据来自清华同方的《中国学术期刊引证报告》、《万方学术期刊引证报告》	基金论文比、总被引频次、影响因子、web 下载率、二次文献转载量	专家定性指标	采用层次分析法确定指标权重	1.首次遴选出“权威期刊”和“核心期刊”； 2.自主研发了“中国学术期刊评价信息系统”； 3.按一级学科和学报类型分类排序，各类别里划分为五个等级	期刊综合得分排名
中信所：中国科技期刊引证报告	选用中国科技核心期刊 1765 种，这些期刊构成了中国科技论文与引文数据库	全部是从该引文数据库统计得出。	<b>期刊引用指标：</b> 总被引频次、影响因子、即年指数、他引率、引用刊数、扩散因子、学科扩散指标、学科影响指标、被引半衰期； <b>期刊计量指标：</b> 来源文献量、文献选出率、参考文献量平均引文数、平均作者数、地区分布数、机构分布数、海外论文比、基金论文比、引用半衰期、离均差率	无	未分配权重，单独统计各项指标的得分	比较全面地给出了各项统计指标，可以用做各项计量研究	未进行任何评价
南京大学：中国人文社会科学期刊学术影响力报告	以 CSSCI 引文数据库为统计源，包括 500 多种刊物	期刊学术规范化指标和学术成果引用指标均来自 CSSCI。	<b>期刊学术规范量化指标：</b> 期刊篇均引用文献数、本学科论文比例、本机构论文比例、期刊作者地区分布、基金论文比例、有机构论文比例； <b>学术成果的引用指标：</b> 期刊被引次数、期刊即年指数、期刊影响因子、期刊被引广度、期刊半衰期； <b>重要二次文献收录指数；Web 即年下载率</b>	无	按照各指标重要性人工分配权重	1.除了常规的引文分析指标外，另外考虑了一些学术规范量化指标； 2.对某些评价指标如被引速率、影响因子、被引广度作了修正； 3.评价角度丰富多样，充分利用数据源从作者、地区、机构等多角度来分析学术影响力； 4.在各项数据	学术规范量化指标、期刊被引次数、被引速率、期刊影响因子、被引广度、二次文献转载等单项指标和

						统计之后,并依次根据数据分析结果附加一定的文字解释数据背后隐藏的事实	综合值分析
名称	选文范围	数据源获取	评价指标		指标权重分配	创新之处	评价角度
			定量	定性			
北京大学:《中文核心期刊要目总览》	中国人文社会科学引文数据库、中国博士学位论文全文数据库、中国重要会议论文全文数据库、中国科学引文数据库、中国期刊全文数据库、全国报刊索引数据库、复印报刊资料全文数据库、中国社会科学文摘、新华文摘、高等学校文科学术文摘、重要报纸转摘文献	有数据库的统计源由编制单位按项目要求统计,书本式的统计源由各学科研究人员手工统计	被索量、被摘量、被引量、他引量、被摘率、影响因子获奖或被重要检索工具收录基金论文比、Web 下载量	针对初选核心期刊,专家再进行评审调整期刊表的排序及内容	征求学科专家意见,确定权重	影响因子、被摘率、基金论文比等公式都作了修改	期刊综合得分排名

	数据库、中国学术期刊文摘						
社 科 院：《中 国人文 社会科 学核心 期刊要 览》	中国人文社会科学引文数据库 (CHSSCD)：收录来源期刊 700 种左右； 《中国社会科学文摘》、《新华 文摘》、《高等学校文科学术文 摘》；重要报纸理论版版和核心 期刊中转摘的文章；复印报刊 资料； 公开发表期刊统计数据	同左	<b>被引率</b> ：期刊学科影响因 子、期刊影响因子、总被 引频次； <b>转摘率</b> ：三大文摘转摘 率、核心报刊转摘率、复 印报刊资料转摘率； <b>借阅率</b> ； <b>下载率</b> ；	专家定性 评价：设 置五项指 标，由专 家对候选 核心期刊 打分	强制判定法：评 价指标两两比 较，两项指标同 等重要打 2 分， 一项比另一项 重要分别打 3 分和 1 分；一项 比另一项更重 要分别打 4 分 和 0 分，然后归 一化处理	1.在对候选核心刊物定性评 价方面设置指标由专家打分； 2.研究分为两大部分：分学科 研制、综合性核心期刊学科引 用分析研究	

从以上各体系之间的对比，可发现：

- 总体上，上述五大文献计量评估体系是围绕“期刊评价”展开的，所设置的评价指标均针对“期刊”设置。
- 五大文献计量评估体系的评价指标，以引文分析指标为主，然后辅助一些期刊计量指标，Web 下载率等网络计量学的指标也得到了应用；其中期刊的被引情况、转载情况、被下载情况、半衰期、被引广度等指标，是各体系均重点采用的核心指标。
- 五大评估体系均偏重定量方面的评价，有三家评价体系中有定性评价指标，并且其定性评价指标过于简单。
- 五大文献计量评估体系评价指标的权重分配，均以人工分配权重（如专家调研）为主，其中武汉大学的评价体系还采用了层次分析法。

从评价的创新趋势可以看到，有概念上的创新，如武汉大学提出的“权威期刊”与“核心期刊”，南京大学的期刊学术规范量化指标；有公式上的重新修订；有新的期刊评价信息系统的诞生；也有深度挖掘数据背后隐藏的重要信息，如南京大学从多角度做出的评价分析。

#### **4.1.4.2 对现有文献计量指标对比分析**

根据以上分析，我们对各评价体系所涉及到的，以及国际上近年新提出的引证计量指标进行了汇总，并确定以下原则作为引证计量指标的遴选原则：1) 可针对单篇论文进行统计分析；2) 能够较为有效地反映论文的影响力；3) 突出反映论文影响力某方面的特征，不与其他指标重合，也不能被其他指标替代；4)

指标数据可以较为容易地在引文数据库中获取。

经过细致分析，对人文社会科学论文的文献计量评估指标做出取舍，见表 16。

表 16 文献计量指标遴选结果

期刊评估指标	分析	是否采用
被引用频次	显示论文被引用的情况； 可针对单篇论文计算	采用“论文被他引”的情况，作为核心指标
影响因子	针对期刊的指标，但针对论文而言，数据量太小	不采用
扩散因子	显示引用论文所覆盖的学科范围或期刊范围	不采用，对于论文而言，按学科计算广度较为复杂，区分度过小
学科影响广度	在所有期刊范围内，引用该刊的期刊数量与其所在学科全部期刊数量之比	
学科扩散因子	期刊所在学科内引用该刊的期刊数占全部期刊数量的比例	
被引半衰期	显示论文的生命力；可针对单篇论文计算，单论文的被引半衰期统计周期长，区分度小	不采用
web 下载率	可针对单篇论文计算；在一定程度反映论文质量	采用，但要关注论文的时间周期
转载情况	可针对单篇论文计算，有效反映论文质量	采用
基金论文比	针对期刊，不能改造为针对单篇论文计量，同行及时评议时更为准确	不采用，已在同行评议指标中体现
即年指标	40%左右的人文社会科学期刊在当年没有被引	不采用

期刊评估指标	分析	是否采用
	用, 数据量过小, 区分度小, 因而更不适合单篇论文的评估	
载文量地区分布	不能有效反映论文的质量	不采用
合著情况	不能有效反映论文的质量	不采用
参考文献数	不能有效反映论文的质量	不采用
H 指数 (高引用次数 high citations)	在一定期间内, 某作者发表的论文至少有 h 篇的被引频次不低于 h 次。针对作者进行评估, 难以用于单篇论文评估, 区分度过小	不采用
g 指数	论文按被引次数排序后相对排前的累积被引至少 $g^2$ 次的最大论文序次 $g$ , 亦即第 $(g+1)$ 序次论文对应的累积引文数将小于 $(g+1)^2$ 。针对作者进行评估, 难以用于单篇论文评估, 区分度过小	不采用

#### 4.1.4.3 结论分析

综合上述对文献计量指标的对比分析, 课题组最终选取了三个与论文质量密切相关的文献计量指标: 被他引频次、被转载次数、Web 下载频次, 具体阐述详见《指标体系》实施方案 2.3.1 部分。

#### 主要参考文献

- (1) Braunt, Glazel. GRUPP H·The scientometriweight of 50 nations in 27 science areas, 1989-1993·Part III life sciences [J]·Scientometrics, 1996, 34: 207-237.

- (2) Donovan,C. A Review of Current Australian and International Practice in Measuring the Quality and Impact of Publicly Funded Research in the Humanities, Arts and Social Sciences[EBOL],REPP Discussion Paper 2005(3), [http://repp.anu.edu.au/papers/2005\\_Setting\\_the\\_Scene.pdf](http://repp.anu.edu.au/papers/2005_Setting_the_Scene.pdf).
- (3) Eglen, P.O. Evaluation of scientists by journal impac[M] //WEINGART P, SEHRINGER R, WINTERHAGER M,Eds-Representations of Science and Technology·Leiden:DSWO Press, 1992: 240-250.
- (4) Eglen, P.O. Why the impact factor of journals should not bused for evaluating research [ J]·British Medical Journa, 1997, 314: 498-502.
- (5) Garfield, E. The history and meaning of the journal impacfactor [J]. JAMA, 2006, 295: 90-93.
- (6) Henkel, Mary. The Modernisation of Research Evaluation:The Case of the UK[J]. Higher Education, 1999(38):105-122.MONASTERSKY R·The number that s devouring scienc(2005-10-14)·[EBOL]·[2009-01-13]·<http://chronicle.com/free/v52/i08/08a01201.htm>.
- (7) Laudel, Grit. Conclave in the Tower of Babel: How Peers Review Interdisciplinary Research Proposals[J], Research Evaluation, 2006-15(1):57-68.
- (8) Laudel, Grit. The‘Quality Myth’:Promoting and Hindering Conditions for Acquiring Research Funds [J]. Higher Education,2006(52):375-403.
- (9) Metrics or Peer Review? Evaluating the 2001 UK Research Assessment Exercise in Political Science [EBOL], REPP working paper, 2007-6-28. [http://repp.anu.edu.au/papers/2007\\_ukresearChassess.pdf](http://repp.anu.edu.au/papers/2007_ukresearChassess.pdf).
- (10) Ren, Quan’ e. Evaluation Index System for Achievements of Humanities & social sciences: an Empirical Analysis of Award-Winning Papers (C) //The 6th International Conference on Scientometrics and University Evaluation.2010(11).
- (11) Senses, Fikret.Difficulties and Trade-offs in Performance Evaluation in Social Sciences-A Turkish Perspective[EBOL].ERC Working Paper In Economic,2003-03-11.ht-tp://www.erc.metu.edu.tr/menu/series03/0311.pdf, 2010-01-02.
- (12) Yan , Erjia; Ying Ding. Weighted Citation: An Indicator of an Article ’ s Prestige[J],Journal of the American Society for Information Science and Technology,2010-61(8):1635-1643.
- (13) Schubert A. Using the h-index for assessing single publications [J]. Scientometrics, 2009, 78 (3): 559-565.
- (14) “欧洲研究区” 网站.What Is the European Research Area [EB / OL] .<http://europa.eu.int/comm/resear/era/index>.
- (15) 《北京大学人文社会科学研究优秀成果评奖办法(试行)》
- (16) 《高等学校人文社会科学研究学术规范 (试行)》
- (17) 《国家社会科学基金项目管理办法》
- (18) 《海南省哲学社会科学规划课题最终成果评估指标体系》
- (19) 《黑龙江省高校人文社会科学研究优秀成果奖励办法实施细则》
- (20) 《湖南省第十届哲学社会科学优秀成果评奖》
- (21) 《吉安市哲学社会科学优秀成果评奖实施办法》
- (22) 教育部《授予博士、硕士学位和培养研究生的学科、专业目录》(1997年)

- (23) 《辽宁省哲学社会科学成果评奖办法》
- (24) 《辽宁省财政科研项目成果评估指标体系》
- (25) 《全国优秀博士学位论文通讯评议评价意见表》
- (26) 《陕西省哲学社会科学优秀成果评奖学科分类办法》
- (27) 《四川大学哲学社会科学研究评价指标体系》
- (28) 《四川省哲学社会科学优秀成果评奖》
- (29) 《云南省第九次(2004 年度)哲学社会科学优秀成果评奖实施细则》
- (30) 《浙江大学社会科学类研究生学术规范》
- (31) 《浙江省社会科学界联合会社会科学研究优秀成果奖评审办法》
- (32) 《浙江省哲学社会科学优秀成果评奖办法实施细则》
- (33) 《中国高校人文社会科学研究优秀成果奖励暂行办法》
- (34) 卜卫, 周海宏, 刘晓红. 社会科学成果价值评估[M]. 社会科学文献出版社, 1999.
- (35) 蔡曙山. 代表性学术成果作为艺术人文与社会科学评估指标的意义——兼论 CSCI 的科学评估功能及其发展方向[J]. 江西社会科学, 2004 (6): 208-220.
- (36) 蔡曙山. 论人文社会科学的科学化、规范化管理——兼析《中华读书报》的不实报道及其错误观点[J]. 学术界, 2001(6).
- (37) 蔡曙山. 中国社会科学发展和社会科学评价[J]. 学术界, 2002 (3).
- (38) 蔡毅. 建立一套良好的学术评价体系[J]. 学术界, 2003(6): 58-72.
- (39) 陈平. 新时期人文社会科学学术评价机制研究[C]. 广州:首届中国高校社会科学科研管理青年论坛, 2009.
- (40) 陈力丹. 谈谈人文社会科学成果的评估标准[J]. 华中科技大学学报(社会科学版), 2003 (1): 14-17.
- (41) 程淑, 桂林, 冀航. 主观评分的归一化算法及误差分析[J]. 高等函授学报(自然科学版). 2007(10): 28-30.
- (42) 褚兆麟. 浅论人大复印报刊资料对社科期刊及所载学术论文的评价和鉴定作用[J]. 社会科学家, 1998 (2) :91-94.
- (43) 邓毅. 社科成果评估指标体系探析[J]. 现代教育论丛, 2004 (6): 21-26.
- (44) 丁军强, 吴桂鸿. 试论社会科学研究成果的评估标准[J]. 科技管理研究, 2007 (6): 87, 91-92.
- (45) 丁军强等. 社会科学研究成果评估指标体系构建之管见[J]. 评估与预测, 2006(4): 113-115.
- (46) 丁学东. 文献计量学基础[M]. 北京大学出版社, 1992.
- (47) 顾海兵. 人文社会科学学术评价规则的时空维度[J]. 学术研究, 2004(10).
- (48) 郭碧坚. 同行评议制: 方法、功能、理论、指标[J]. 科学学研究, 1994, (03) .
- (49) 郭开选. 文科学术论文学术价值的估量[J]. 渭南师范学院学报, 2002 (7): 90-92.
- (50) 郭丽芳. 评价论文学术质量的文献计量学指标探讨[J]. 现代情报, 2005 (3) :11-12.
- (51) 国家社会科学院外事局. 美国人文社会科学现状与发展[M]. 社会科学文献出版社, 2001: 370-391.
- (52) 贺忠德. 社会科学研究成果鉴定标准和评估方法初探[J]. 社科论坛, 1998 (1-2): 66-11.
- (53) 胡红生. 人文社会科学成果的评估标准[J]. 武汉理工大学学报(社会科学版), 2001 (8): 316-319.
- (54) 胡明铭, 黄菊芳. 同行评议研究综述[J]. 中国科学基金, 2005(04).
- (55) 胡显章、杜祖贻、曾国屏. 国家创新系统与学术评价制度[M]. 山东: 山东教育出

- 版社, 2000.
- (56) 贾岩. 编辑学视角中的社科论文评估[J]. 湖南大众传媒职业技术学院学报, 2006 (1): 82-83.
- (57) 金碧辉,汪寿阳,汪冰等. 国际论文与国内论文合一计算方法研究[J]. 管理科学学报, 1999, 2 (3): 59-65.
- (58) 姜晓辉等编. 中国人文社会科学核心期刊要览(2008年版)[M]. 北京: 社会科学文献出版社, 2009.
- (59) 金武刚. 当前我国社会科学研究的定量分析—兼论社会科学研究成果评价体系的构建[D]. 华东师范大学, 2001.
- (60) 雷二庆. 单篇科技期刊论文评估方法初探[C]. 第十届全国科技评价学术研讨会精选论文集, 2010.
- (61) 廖萍. 高校社会科学成果评估指标体系研究[J]. 科技管理研究, 2008(5): 169-170.
- (62) 李存娜. 人文社会科学评价的问题与反思[J]. 学术界, 2004(2): 154-167.
- (63) 李宏伟. 成果评价的原则与日美评价的比较研究[J]. 自然辩证法研究, 1996(2): 38-42.
- (64) 李志平. 高校社会科学成果的一种四维评价方法初探[J]. 科技管理研究, 2004(1): 89-91.
- (65) 刘大椿. 中国人文社会科学评估问题之审视[J]. 重庆大学学报(社会科学版), 2009 (1): 54-59.
- (66) 刘大椿等. 人文社会科学研究成果评估体系研究[M]. 经济科学出版社, 2009.9.
- (67) 刘莉. 欧洲各国大学科研评价及其启示[J]. 科学学与科学技术管理, 2005(2): 86-90.
- (68) 刘明. 学术评估制度批判[M]. 长江文艺出版社, 2006.
- (69) 刘益东. 创新力度: 评价学术成果的首要标准[J]. 科技创新导报, 2009(36): 107-109.
- (70) 刘益东. 试论超越同行评议的复合型学术评估法[J]. 自然辩证法研究, 2004(1): 98-102.
- (71) 娄群策. 社会科学评价的文献计量理论与方法[M]. 华中师范大学出版社, 1999.
- (72) 卢渝. 社会科学研究成果分类标准及其量化途径[J]. 社会科学, 1992(4): 73-76.
- (73) 罗平安. 成果、论文水平评估标准. 非出版物.
- (74) 么大中. 关于社会科学优秀成果评选标准问题[J]. 科研管理, 1986(1): 63-66.
- (75) 邱均平等编. 中国学术期刊评价研究报告: RCCSE 权威、核心期刊排行榜与指南[M]. 北京: 科学出版社, 2009.
- (76) 庞学铨. 论人文科学的价值与功能[J]. 中共浙江省委党校学报, 2009(2): 11-17.
- (77) 裴长洪. 美国人文社会科学现状与发展[J]. 社会科学文献出版社, 2001: 371-375.
- (78) 任全娥. 人文社会科学研究成果评价指标体系研究[J]. 大学图书馆学报, 2009(5): 51-55.
- (79) 沙似鹏, 郑礼, 郭才伯, 张毅. 人文、社会科学研究成果评估指标体系初探[J]. 上海高教研究, 1994(1): 74-78.
- (80) 苏为华. 多指标综合评价理论与方法研究[M]. 中国物价出版社 2001.
- (81) 苏新宁. 中国人文社会科学学术影响力报告(2000-2004)[M]. 中国社会科学出版社, 2007.
- (82) 苏新宁. 中国人文社会科学学术影响力报告(2009)[M]. 中国社会科学出版社, 2009.
- (83) 孙浩等. 论建立社会科学成果评价指标体系[J]. 河北师范大学学报:自然科学版, 1998, 22(2): 139-144.
- (84) 谭春晖. 国外对人文社会科学研究成果的评价[J]. 科学管理研究, 2006(4).

- (85) 谭贤楚, 刘伦文, 叶明. 人文社会科学成果评估: 必要性、理论及方法——基于社会学的分析与思考[J]. 湖北民族学院学报(哲学社会科学版), 2006(2): 155-158.
- (86) 唐德章, 夏元林. 社会科学研究成果定量评估方法初探[J]. 社会科学研究, 1989(1): 78-87.
- (87) 王兰敬, 杜慧平. 欧美人文社会科学评价的现状与反思[J]. 南京大学学报(哲学·人文科学·社会科学), 2010(1): 111-118.
- (88) 王 瑜. 我国高校人文社会科学研究成果评价[J]. 科技管理研究, 2008(4): 83-89.
- (89) 王贯中. 高校科研成果评估指标体系的构建[J]. 科技促进发展, 2008(8): 43-46.
- (90) 王晓丽. 浅议高校人文社科研究成果的有效评价[J]. 科技管理研究, 2005(6): 112-114.
- (91) 韦莉莉, 孙晶. 人文社会科学评价理论与实践评析[J]. 社会科学管理与评论, 2006(4).
- (92) 文庭孝, 邱均平. 论科学评价理论研究的发展趋势[J]. 科学学研究, 2007(2): 204-209.
- (93) 吴桂鸿. 社会科学研究成果评估指标体系研究[D]. 湖南大学, 2006.
- (94) 吴平. 社会科学评价程序与方法运用[J]. 情报科学, 2006(6): 814-818.
- (95) 吴树青. 创新是繁荣哲学社会科学的灵魂[N]. 光明日报, 2006-01-16.
- (96) 熊树民. 日本的基础研究成果评价[M]. 世界科技研究与发展, 1997: 84-85.
- (97) 许树柏. 层次分析法原理[M]. 天津大学出版社, 1988: 89.
- (98) 杨家栋, 秦兴方. 社会科学研究成果的评估及其指标体系[J]. 齐鲁学刊, 2001(2): 122-128.
- (99) 杨霞, 熊世春. 社会科学研究成果综合评估指标体系研究[J]. 山西师大学报(社会科学版), 2008(5): 51-54.
- (100) 杨育华. 试论社会科学研究成果的评估[J]. 宁夏社会科学, 2002(9): 125-128.
- (101) 叶继元. 中国哲学社会科学学术期刊合理布局研究[M]. 社会科学文献出版社, 2008.
- (102) 叶继元. 学术期刊的定性与定量评价[J]. 图书馆论, 2006, 26(6): 54-58.
- (103) 叶蓬. 人文社会科学研究成果评估指标体系分析[J]. 探求, 2001(1): 60-67.
- (104) 袁培国. 多指标加权值对多科学学术期刊排序初探[J]. 重庆大学学报(社会科学版), 2009(15): 74-82.
- (105) 喻承久, 张勇. 社会科学成果评估指标体系分析[J]. 空军雷达学院学报, 2005(9): 73-75.
- (106) 张慧颖, 张卫滨, 张颖春. 哲学社会科学学术成果评价方法的比较研究[J]. 理论与现代化, 2007(1): 108-113.
- (107) 中国社会科学院国际合作局. 英国人文社会科学研究的评价体系 [EBOL]. <http://bis.cass.cn>, 2007-03-21.
- (108) 周克兴. 人文社会科学与国家创新体系[J]. 科学学与科学技术管理, 2000(7): 34.
- (109) 周来祥. 人文社会科学研究的特点与规律[J]. 文史哲, 2003(1): 5-7.
- (110) 周露阳. 学术论文的创新途径: 基于创新因素的分析[J]. 科技管理研究, 2006(8): 195-197.
- (111) 朱强、戴龙基、蔡蓉华等编. 中文核心期刊要目总览(2008年版)[M]. 北京: 北京大学出版社, 2008.
- (112) 朱少强. 国内外人文社会科学研究评价综述[J]. 评价与管理, 2007.12, 39-73.
- (113) 朱少强. 论科学建制背景下的人文社会科学研究评价[J]. 评价与管理, 2008.12, 44-47.

## 4.2

# 人文社会科学论文质量评估指标体系 问卷调查报告

中国人民大学人文社会科学学术成果评价研究中心

2010年11月15日

## 4.2.1 问卷调查概述

### 4.2.1.1 调查背景

本报告是“人文社会科学论文质量评估指标体系”设计的分报告之一。截至2010年7月，课题组通过前期全面深入的文献调研，系统地汇总梳理了“人文社会科学论文质量评估基本指标”。这些指标包括同行评议指标和文献计量指标两大类，其中同行评议指标是指标体系的主体部分。

以同行评议为主的人文社会科学论文质量评估的难点，是如何对论文进行科学分类、设定哪些指标、指标权重如何赋值，以及如何更具可操作性等，这些问题都是当下学术理论界、期刊出版界和科研管理界极为关注的热点问题。关于同行评议评估指标体系的设计，学术评价领域普遍认同的原则是要“从学术共同体的共识性标准出发”。为此，课题组针对论文的分类评估方案，以及同行评议指标的必要程度和权重设置问题，设计了3个调查问卷（见附件调查问卷），并于2010年7月26日至2010年10月26日期间，主要向人文社会科学领域的专家学者、审稿编辑和科研管理者三大群体开展了较大规模的问卷调查。

### 4.2.1.2 调查目的

人们普遍认为，遴选科研成果评估同行评议指标，应体现学术共同体的“共识性判断”标准。为此，我们列出了以下同行评议的关键问题，拟通过问卷调查了解人文社会科学领域专家的共识性看法，以提高同行评议的科学性、有效性和可行性。

- 对人文社会科学论文进行评估，应采取什么样相对科学合理的分类方案？
- 课题组前期遴选的人文社会科学论文同行评议指标，是否是学术共同体普

遍认同的必要指标?

- 课题组前期遴选的人文社会科学论文同行评议指标, 相对重要性程度应怎样处理和分配?
- 人文社会科学各学科专家, 对论文分类、指标设置、指标权重分配的认识, 是否存在明显差异以及有哪些差异?

#### 4.2.1.3 样本分析

本项问卷调查历经 3 个月, 主要通过纸质调查问卷邮寄、电子邮件发送或回收问卷、专家面访等方式开展调查工作。本次调查发放问卷近 800 份, 收回问卷 388 份, 有效问卷 386 份。

##### (1) 受访对象由编辑、学者和科研管理者三类群体组成

正确选择专家群体是问卷调查成败的关键。本次调查主要针对人文社会科学领域的 3 类受访者进行调查分析, 即编辑、学者和科研管理者, 其中前两类受访对象是重点。3 类受访者的样本条件及数量和职业类型分布比例分别见表 17 和图 7。

表 17 受访者的样本条件与数量

受访对象类型	样本特征与条件	样本量 <sup>88</sup>
学术期刊杂志编辑	<ul style="list-style-type: none"> <li>➤ 主要包括期刊杂志的主编、副主编和资深编辑。</li> <li>➤ 这一组专家长期关注和阅读大量的人文社会科学论文, 对论文质量的优劣区分具有丰富的经验。</li> </ul>	109
教学研究的专家学者	<ul style="list-style-type: none"> <li>➤ 在我国各地的高等院校和科研院所的研究学者中随机遴选, 遴选原则是: 第一, 具有副高和正高职称的专家学者; 第二, 在某学科领域中具有较高质量水平的代表性论文或学术专著。</li> <li>➤ 这一组专家长期关注和研究各自学科领域的前沿问题, 能够较好地把握专业性学术论文质量的内涵。</li> </ul>	280
科研管理者	<ul style="list-style-type: none"> <li>➤ 在我国各地的高等院校和科研院所的资深科研管理者中随机遴选。</li> <li>➤ 这一组专家长期工作于科研成果评估的第一线, 对于指标体系的合理性和可操作性具有丰富的经验。</li> </ul>	29

注: 受访者存在职业类型多选情况

<sup>88</sup> 三类受访对象有重合, 统计时按照学者、编辑、科研管理者的优先顺序, 只统计第一职业类型。

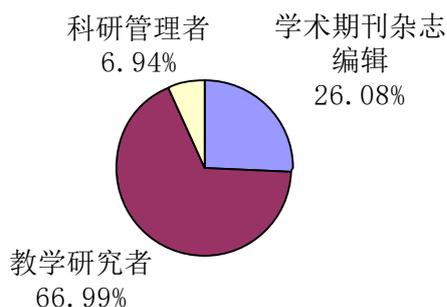


图7 受访对象职业类型分布

## (2) 受访对象覆盖了人文社会科学的主要学科

为保证问卷调查结果能够较好地反映人文社会科学各学科领域的实际情况,我们还对受访专家的学科构成进行了比例控制。以教育部《授予博士、硕士学位和培养研究生的学科、专业目录》(1997年颁布)所确定的人文社会科学一级学科为依据,根据学科规模估计的大小分发问卷数量。本次受访对象共覆盖以下学科(见表18):

表18 受访者的学科分布

一级学科	问卷数	百分比	一级学科	问卷数	百分比
哲学	65/42	10.90%	中国语言文学	60/28	7.30%
理论经济学	40/27	7.00%	外国语言文学		
应用经济学	45/23	6.00%	新闻传播学	40/18	4.70%
法学	65/22	5.70%	艺术学	40/17	4.40%
政治学	70/39	10.10%	历史学	55/35	9.10%
社会学	50/23	6.00%	管理科学与工程	75//34	8.80%
民族学			工商管理		
马克思主义理论	40/10	2.60%	农林经济管理		
教育学	50/31	8.00%	公共管理		
心理学	30/8	2.10%	图书馆、情报与档案管理	45/22	5.70%
体育学	30/7	1.80%	<b>合计</b>	800/386	100%

### 【说明】

- 问卷数:前一个数值指实际分发的问卷数量,后一个指有效回收的问卷数量。
- 百分比:是指该学科有效问卷占分发问卷总量的比例。
- 在统计时,问卷数较少且研究领域相似的部分一级学科的问卷进行了合并;对于填写了两个以上学科的受访者,只选取了所填写的第一学科进行统计。

## 4.2.2 论文分类评估方案

论文分类评估方案是构建科学可行论文质量评估指标体系的基础和前提，至关重要。为此，课题组围绕有代表性的四个分类方案，开展了问卷调查（见附件问卷一）：

方案一：按学科大类划分为人文学科论文和社会科学论文。

方案二：按论文的论述体裁划分为研究论文和综述文章（含述评）。

方案三：按科研活动的性质划分为基础理论研究论文和应用实证研究论文。

方案四：构建一个通用的评价指标集合，根据不同学科（政治学、经济学、管理学、教育学、历史学、哲学、文学等）对指标赋予不同权重。

【说明】每位受访者可同时选择自己认为可行的两个及以上方案。

图 8 展示了四个方案第一选择和第二选择的总体情况。

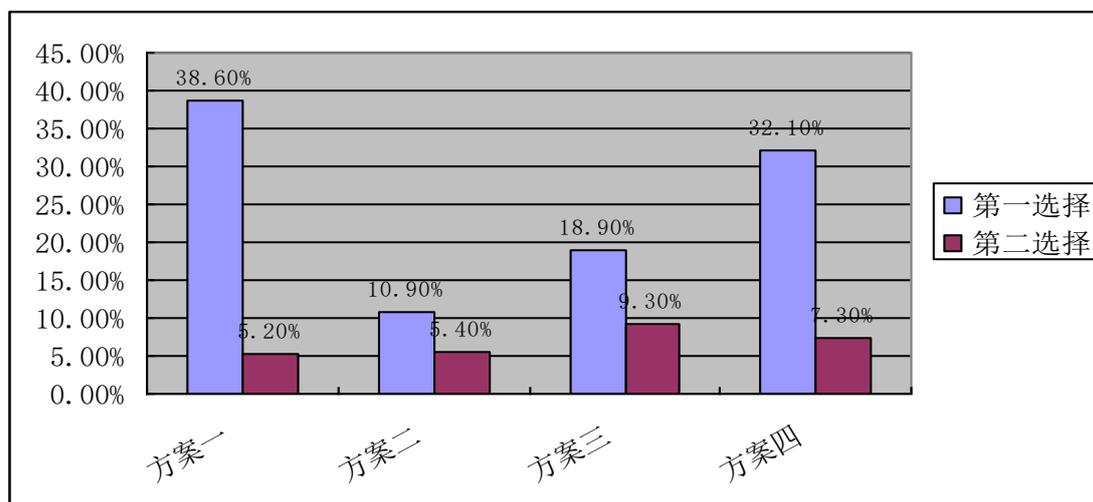


图 8 全体受访者选择分类方案的调查  
(有效问卷 386 份)

#### 4.2.2.1 受访者赞同按“学科分类”方案的比例最高

方案一和方案四两个分类方案，都遵循“按学科分类评估”的基本原则，只是方案四比方案一分类的颗粒度更小。根据图 2，这两类分类方案得到的认同度最高，约 70%。受访者认为这两个分类方案更为合理的主要理由包括：1) 学科之间具有天然的差异，区分度大、可操作性强；2) 同学科内的论文具有可比，也便于找到同行专家，可从学科发展的整体角度判断论文的质量；3) 按学科分类的体系较为成熟系统，符合大多数研究者的习惯，便于检索；4) 相比之下，方案二之所

以不宜采用，是因为“综述文章”在所有论文中所占的比重很小，其它研究论文的差异性得不到充分体现；方案三之所以不宜采用，是因为人文社会科学论文的基础研究和应用研究的性质较难区分，不易操作。

#### 4.2.2.2 受访者赞同“按论述体裁分类”和“按研究性质分类”方案的比例不高

根据图 1，方案二和方案三得到赞同的比例均不高。选择这两类分类方案的主要理由包括：1) 学科交叉趋势日益增强，按学科分类难以区分所有论文；2) 综述类论文的研究方法和论述体裁明显有别于一般研究论文，便于区分操作、包容性大；3) “按研究性质分类”符合管理，有利于明确研究的目标，并且相对来讲更为科学合理、操作性强。

#### 4.2.2.3 部分受访者建议应综合考虑不同分类方案的优势

少部分学者还指出，单一的分类方案很难获取精确的评估数据，应综合考虑论文多个维度属性的特征，进行多维分类：所在学科、论述体裁、研究性质等，都是一篇论文同时需要考虑的属性，尤其是前两者，差异性十分明显，只考虑一种分类方案，难以达到准确；方案一和方案四两种分类，在学科颗粒度上有差异；分类不宜过粗，否则不能体现其重要差异性，也不宜过细，否则难以操作。因此，需要全面进行综合平衡和处置。

#### 4.2.2.4 不同职业类型受访者选择分类方案的偏好差异不明显

如图 9 所示，不同职业类型的受访者，对于分类方案的倾向有一定差异：教学研究群体对各类分类方案的赞同比例相对均衡，其中方案一得到的赞同比例最高；科研管理者更为倾向方案一，赞成方案四的比例相对较低；编辑群体更为赞成方案四，对方案二的赞同比例较低。

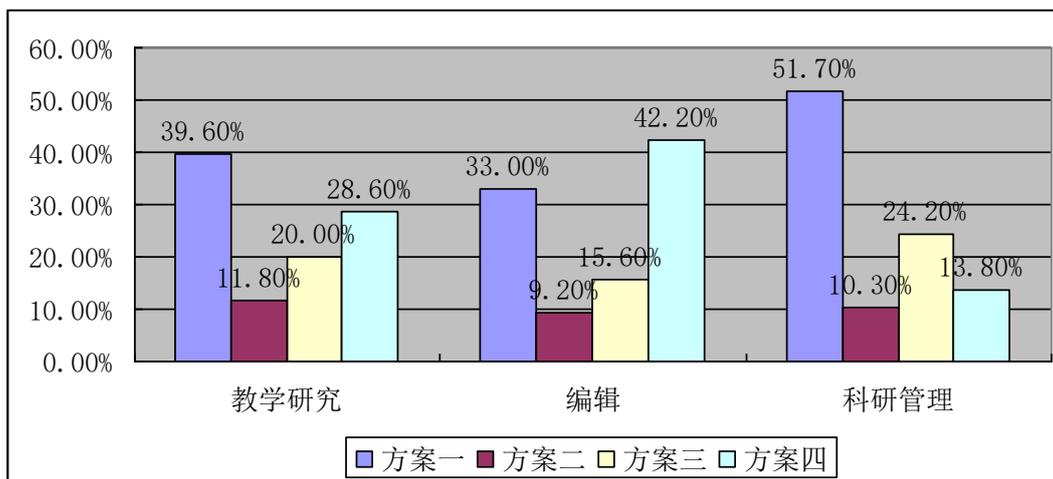


图9 不同职业类型受访者分类方案（第一选择）的调查  
(有效问卷 386 份)

#### 4.2.2.5 不同学科的受访者选择分类方案的偏好差异较为明显

不同学科的受访者在选择分类方案时，呈现出较为明显的差异性，表现在：

- **马克思主义学科的受访者**，无人选择方案二或者方案三；
- **教育学受访者**中把方案四作为第一选择达约到 74.2%，选择方案二或者方案三的百分比之和不超过 7%。
- **历史学和艺术学受访者**中的 60%以上，把方案一作为第一选择。

#### 4.2.2.6 结论分析

- 按学科分类评估是同行评议指标体系应采取的相对较为科学合理可行的分类方案。
- 评估指标体系有必要在权衡科学合理性和可行性的前提下，综合考虑和权衡多种分类方案的优势。

### 4.2.3 论文评价指标的必要程度

指标对考察论文学术质量的必要程度，是决定指标是否应纳入体系的重要依据。

我们在问卷中预先设置了四个必要程度选项：1.很必要，2.有必要，3.必要程度不

大, 4.不必要 (见附件问卷二), 用来了解受访者对指标必要程度的认识。

#### 4.2.3.1 受访者对评估指标必要程度的总体认识基本一致

##### (1) 受访者普遍认为四个“主要指标”对于考察论文质量“有必要”

表 19 全体受访者对主要指标必要程度的看法  
(有效问卷 304 份)

评估指标	评估内容	很必要	有必要	总比例
学术创新程度	1.提出新的(或修正完善已有的)学说、理论、观点、问题、阐释等,对学术发展的促进程度	68.40%	20.00%	88.40%
	2.提出新的(或改进运用已有的)方法、视角等,对学术发展的促进程度	60.20%	29.90%	90.10%
	3.发现的新资料、史料、证据、数据等,对学术发展的促进程度	52.10%	37.60%	89.70%
	4.对已有成果做出新的概括、评析,对学术发展的促进程度	38.20%	45.70%	83.90%
论证完备程度	5.理论前提的科学性	68.00%	23.10%	91.10%
	6.研究方法的有效性	63.60%	30.80%	94.40%
	7.论据的充分可靠性	76.20%	20.90%	97.10%
	8.逻辑的严密系统性	66.70%	28.70%	95.40%
	9.学术引证的规范性	70.40%	26.00%	96.40%
难易程度	10.论题的复杂程度(理论难点或实证研究难度)	26.90%	48.50%	75.40%
	11.资料搜集或处理的难度	24.00%	54.70%	78.70%
社会价值	12.对解决经济、政治、社会建设中问题的推动作用	33.40%	47.00%	80.40%
	13.对思想道德文化建设的促进作用	19.30%	47.80%	67.10%

##### (2) 对“辅助指标”必要程度的看法,受访者争议较大

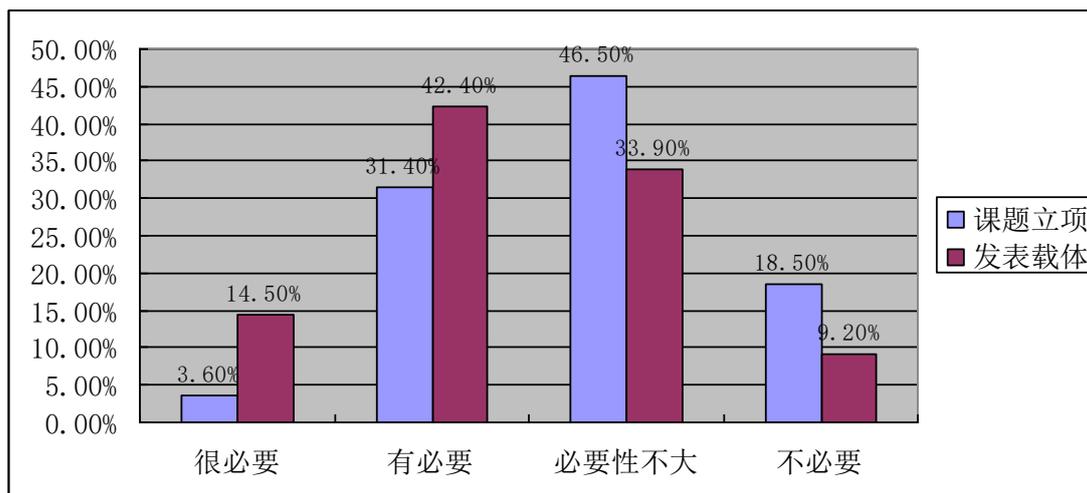


图 10 受访者对“辅助指标”必要程度的看法

(有效问卷 304 份)

如图 10 所示, 56.9%的受访者认为“发表载体”指标“有必要”或者“很必要”。认为“课题立项”指标必要的受访者占 35%; 认为不必要的受访者占 65%。认为“课题立项”指标和“发表载体”指标“必要程度不大”或“不必要”, 主要理由包括: 1) “辅助指标”与论文质量无直接联系, 应当取消; 2) 大多数发表载体与论文的优劣并无因果关系; 3) 一篇论文的优劣与课题立项之间并无因果关系; 4) 创新和社会价值是最重要的指标, 是否有课题立项支持或发表于重要期刊载体不是关键影响因素; 5) 课题立项有明显“官本位”倾向, 不能看得过重。因此, 辅助指标是否应纳入指标体系, 需要进一步进行实证检验。

### (3) 评估指标的各项评估内容在必要程度上呈现差异

根据表 19 所示的结果, 结合判断必要程度的百分比的比例和均值两项数据, 可将各指标的评估内容的必要程度做如下排列 (从高到低): 7-1-5-2-6-8-3-9-4-12-10-11-13-15-14 (14 为课题立项, 15 为发表载体)。此外, 以下两点值得关注:

1) “对已有成果做出新的概括、评析”的必要程度明显低于其他创新内容。

如表 5 所示，“学术创新程度”指标的评估内容中，认为前三项内容“很必要”的受访者数均超过 50%，远远多于认为第四项内容——“对已有成果做出新的概括、评析，对学术发展的促进程度”（见附件问卷二）。第四项内容只有 38.2%的学者认为“很必要”；45.7%的受访者认为“有必要”。由于此项内容针对“综述文章”，由此可见，对“综述文章”学术创新程度评估的必要程度，低于一般研究论文。

## 2)“对思想道德文化建设的促进作用”的必要程度明显低于其他“社会价值”。

如表 5 所示，“社会价值”指标的评估内容中，“对思想道德文化建设的促进作用”这项评估内容，只有 19.3%的学者认为是“很必要”，47.8%的学者认为是“有必要”；而“社会价值”指标的另一项评估内容“对解决经济、政治、社会建设中问题的推动作用”，则有 33.4%的受访者认为“很必要”，47.0%的人认为“有必要”。这说明“社会价值”指标对评价论文质量“有必要”并非“很必要”；“对思想道德文化建设的促进作用”这项评估内容的必要程度低于另一项评估内容。

### 4.2.3.2 受访者对个别评估指标必要程度的认识存在差异

#### (1)“人文学科论文”与“社会科学论文”指标必要程度存在微小差异

表 20 “人文学科论文”与“社会科学论文”指标必要程度对比  
(有效问卷 162 份)

指标	评估内容	很必要		有必要	
		人文	社科	人文	社科
学术创新程度	提出新的(或修正完善已有的)学说、理论、观点、问题、阐释等,对学术发展的促进程度	84.60%	79.40%	13.60%	20.00%
	提出新的(或改进运用已有的)方法、视角等,对学术发展的促进程度	65.20%	68.80%	28.00%	29.40%
	发现的新资料、史料、证据、数据等,对学术发展的促进程度	62.70%	53.50%	32.30%	42.10%
	对已有成果做出新的概括、评析,对学术发	41.50%	42.10%	44.00%	41.50%

	展的促进程度				
<b>论证完备程度</b>	理论前提的科学性	68.10%	81.10%	21.50%	15.10%
	研究方法的有效性	57.80%	71.90%	34.80%	25.00%
	论据的充分可靠性	72.70%	80.60%	23.60%	18.10%
	逻辑的严密系统性	69.60%	78.10%	24.20%	20.00%
	学术引证的规范性	70.60%	72.30%	25.80%	25.20%
<b>难易程度</b>	论题的复杂程度（理论难点或实证研究难度）	36.30%	37.30%	39.40%	43.70%
	资料搜集或处理的难度	35.00%	33.50%	43.80%	50.00%
<b>社会价值</b>	对解决经济、政治、社会建设中问题的推动作用	27.30%	52.20%	41.00%	39.00%
	对思想道德文化建设的促进作用	39.10%	23.30%	36.60%	45.30%
<b>课题立项</b>	如国家社科规划课题，省部级重点课题等	11.30%	10.10%	26.30%	32.70%
<b>发表载体</b>	如 SCI、SSCI、A&HCI 来源刊、光明日报、人民日报，南大、北大、社科院的核心期刊	14.90%	14.40%	43.50%	45.00%

从表 20 可以看出，赞同将论文分为人文学科与社会科学的受访者，认为两个学科大类论文的评估指标，在必要程度上应有所差异，但差异不大，主要体现在：

- “学术创新程度”：人文学科论文在“提出新的（或修正完善已有的）学说、理论、观点、问题、阐释”和“发现的新资料、史料、证据、数据”方面，必要程度高于社会科学论文。
- “论证完备程度”：社会科学论文对该项指标的各项评估内容的必要程度（除学术规范性外），均明显高于人文学科论文。
- “社会价值”：人文学科论文更强调“对思想道德文化建设的促进作用”；社会科学论文更强调“对解决经济、政治、社会建设中问题的推动作用”。

“难易程度”、“课题立项”、“发表载体”：这三个指标的必要程度在两类论文对比中差异都不大。

**(2) “综述文章”与“研究论文”的评估指标在必要程度上存在显著差异**表 21 “综述文章”和“研究论文”指标必要程度对比  
(有效问卷 72 份)

评估指标	评估内容	很必要		有必要	
		综述	研究	综述	研究
学术创新程度	提出新的(或修正完善已有的)学说、理论、观点、问题、阐释等,对学术发展的促进程度	10.00%	90.50%	31.70%	7.90%
	提出新的(或改进运用已有的)方法、视角等,对学术发展的促进程度	13.60%	74.60%	37.30%	23.80%
	发现的新资料、史料、证据、数据等,对学术发展的促进程度	30.50%	63.50%	40.70%	34.90%
	对已有成果做出新的概括、评析,对学术发展的促进程度	75.80%	41.90%	17.70%	37.10%
论证完备程度	理论前提的科学性	33.30%	76.20%	34.90%	19.00%
	研究方法的有效性	31.70%	77.80%	40.00%	17.50%
	论据的充分可靠性	53.30%	88.90%	35.00%	11.10%
	逻辑的严密系统性	50.00%	87.30%	36.70%	11.10%
	学术引证的规范性	73.80%	76.20%	19.70%	19.00%
难易程度	论题的复杂程度(理论难点或实证研究难度)	18.00%	38.10%	36.10%	49.20%
	资料搜集或处理的难度	35.00%	30.20%	35.00%	47.60%
社会价值	对解决经济、政治、社会建设中问题的推动作用	15.00%	28.60%	48.30%	57.10%
	对思想道德文化建设的促进作用	11.70%	20.60%	40.00%	49.20%
辅助指标	课题立项 如国家社科规划课题,省部级重点课题等	3.30%	9.50%	25.00%	31.70%
	发表载体 如 SCI、SSCI、A&HCI 来源刊、光明日报、人民日报,南大、北大、社科院的核心期刊	8.30%	14.30%	35.00%	44.40%

从表 21 可以看出,赞同将论文分为“综述文章”与“研究论文”的受访者,认为两类论文的评估指标,在必要程度上差异不大,主要体现在:

- “学术创新程度”:“综述文章”的创新着重体现在“对已有成果做出新的概括、评析”上,其他三项创新内容的必要程度,均明显低于“研究论文”。

- “论证完备程度”：“综述文章”对该项指标的各项评估内容的必要程度（除学术规范性外），均明显低于“研究论文”。
- “难易程度”：对“论题的复杂程度”的评估必要程度，“综述文章”明显低于“研究论文”。
- “社会价值”、“课题立项”、“发表载体”：“综述文章”的指标必要程度明显低于“研究论文”。

### (3) “基础理论研究论文”与“应用实证研究论文”部分指标的必要性存在一定差异

表 22 “基础理论研究论文”和“应用实证研究论文”指标必要程度对比  
(有效问卷 116 份)

评估指标	评估内容	非常必要		有必要	
		基础	应用	基础	应用
学术创新程度	提出新的(或修正完善已有的)学说、理论、观点、问题、阐释等,对学术发展的促进程度	86.00%	24.30%	11.20%	45.80%
	提出新的(或改进运用已有的)方法、视角等,对学术发展的促进程度	62.00%	46.20%	32.40%	41.50%
	发现的新资料、史料、证据、数据等,对学术发展的促进程度	52.30%	53.30%	40.40%	26.70%
	对已有成果做出新的概括、评析,对学术发展的促进程度	38.00%	22.90%	45.40%	45.70%
论证完备程度	理论前提的科学性	84.30%	57.70%	12.00%	30.80%
	研究方法的有效性	59.60%	74.00%	35.80%	22.10%
	论据的充分可靠性	73.40%	80.80%	24.80%	19.20%
	逻辑的严密系统性	81.70%	57.10%	15.60%	34.30%
	学术引证的规范性	74.30%	56.20%	21.10%	32.40%
难易程度	论题的复杂程度(理论难点或实证研究难度)	34.60%	26.90%	43.90%	48.10%
	资料搜集或处理的难度	25.00%	36.90%	51.90%	43.70%
社会价值	对解决经济、政治、社会建设中问题的推动作用	28.30%	68.60%	48.10%	25.70%
	对思想道德文化建设的促进作用	23.60%	26.00%	43.40%	48.10%
辅助	课题立项 如国家社科规划课题,省部级重点课题等	6.50%	10.50%	33.30%	31.40%

指标	发表载体	如 SCI、SSCI、A&HCI 来源刊、光明日报、人民日报, 南大、北大、社科院的核心期刊	19.30%	16.00%	44.00%	37.70%
----	------	--	--------	--------	--------	--------

从表 22, 可以看出, 赞同将论文分为“基础理论研究论文”与“应用实证研究论文”的受访者, 认为两类论文的评估指标, 在必要程度上差异不大, 主要体现在:

- “学术创新程度”: 除了“发现的新资料、史料、证据、数据等”一项以外, 其他三项创新内容的必要程度, “基础理论研究论文”均明显高于“应用实证研究论文”。
- “论证完备程度”: “基础理论研究论文”更为强调“理论前提的科学性”、“逻辑的严密系统性”和“学术引证的规范性”; “应用实证研究论文”则更为强调“研究方法的有效性”和“论据的充分可靠性”。
- “难易程度”: 对“基础理论研究论文”的考察, 更为强调“论题复杂程度”; 对“应用实证研究论文”的考察, 更为强调“资料搜集或处理的难度”。
- “课题立项”、“发表载体”: 两类论文评估指标的必要程度差异不大。

#### 4.2.3.3 结论分析

根据问卷调查结果, 我们可以得出认为: 尽管各指标及其评估内容, 在必要程度上存在一定差异, 但可初步认为调查的指标的必要程度普遍较高。具体来讲:

- “学术创新程度”和“论证完备程度”两项指标的均得到了 80%以上的学者赞同, 因此是“非常必要”的指标;
- “难易程度”和“社会价值”、“发表层次”三个指标, 大部分学者认为“有必要”, 但不是“非常必要”;

- “课题立项”指标约 35%的学者认为“有必要”，因此，该指标有待进一步考察。
- “综述文章”与“研究论文”在评估指标的 necessary 程度上差异较为显著，在指标体系的分类方案和指标设置上，应予以充分重视。

#### 4.2.4 论文评价指标的权重分配

各项指标权重的分配，对评估结果将产生重要影响。为使权重的调查结果更为精确，经文献调研结果和综合对比分析，本部分采用“层次分析法”，构建“两两比较的矩阵”调查各项指标的相对重要程度（见附件问卷三）；之后将矩阵利用 MATLAB 软件进行检验计算；最后利用 SPSS 软件进行统计分析 and 处理。

经过检验，问卷三的有效问卷数量共 176 份，有效矩阵数量为 301 个（含分类方案多选情况）。权重值按六项指标权重和为 100% 计算。

##### 4.2.4.1 “主要指标”和“辅助指标”权重分配比例约为“85%：15%”

主要指标即“学术创新程度”、“论证完备程度”、“难易程度”和“社会价值”，这四个指标用于同行评议定性判断的主体指标；辅助指标即“课题立项”和“发表载体”两个用于补充同行评议判断的指标。

若不区分学科和论文类型（即选择分类方案四），共有 65 个矩阵有效。我们对这 65 个矩阵进行了综合的统计分析，发现“主要指标”和“辅助指标”的比例如表 23 所示。

表 23 主要指标和辅助指标权重问卷调查结果  
(有效问卷 65 份)

指标类别	权重值
主要指标	84.76%
辅助指标	15.24%

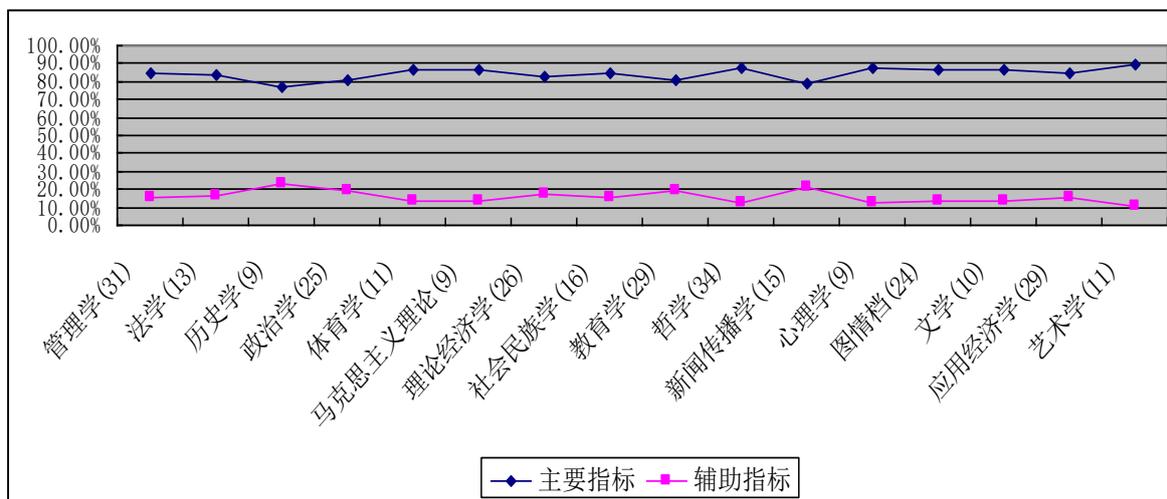


图 11 各学科“主要指标”和“辅助指标”权重分配对比  
(注：括号内数字为“有效矩阵数量”)

如图 11 所示，不同学科背景的受访者对指标权重分配的看法存在明显差异：“艺术学”背景受访者所认为的“主要指标”权重最高、“辅助指标”权重最低，其比例为“89.51%：10.48%”；历史学背景受访者所认为的“主要指标”权重最低、“辅助指标”权重最高，其比例为“77.36%：22.65%”。

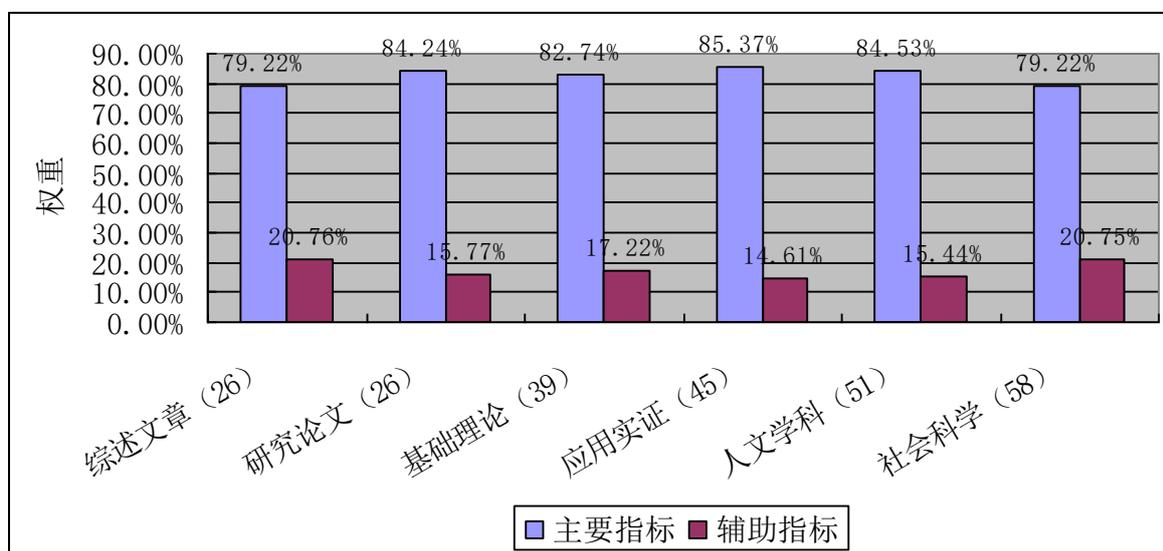


图 12 不同类型论文权重分配对比  
(注：括号内数字为“有效矩阵数量”)

如图 12 所示，选择不同分类方案的受访者进行统计，发现“应用实证研究论

文”的“主要指标”权重最高、“辅助指标”权重最低,其比例为“85.37%:14.61%”;“综述文章”的“主要指标”权重最低、“辅助指标”权重最高,其比例为“79.22%:20.76%”。

#### 4.2.4.2 “人文学科论文”与“社会科学论文”的指标权重分配存在微小差异

如图 13 所示,对选择方案一的受访者进行统计分析发现,“人文学科论文”和“社会科学论文”在指标权重分配上有如下差异:

- “人文学科论文”比“社会科学论文”更偏重“主要指标”,但差异不明显;
- “人文学科论文”的“学术创新程度”、“论证完备程度”和“难易程度”三项指标的权重都略高于“社会科学论文”;“社会科学论文”的“社会价值”指标的权重稍高于“人文学科论文”。

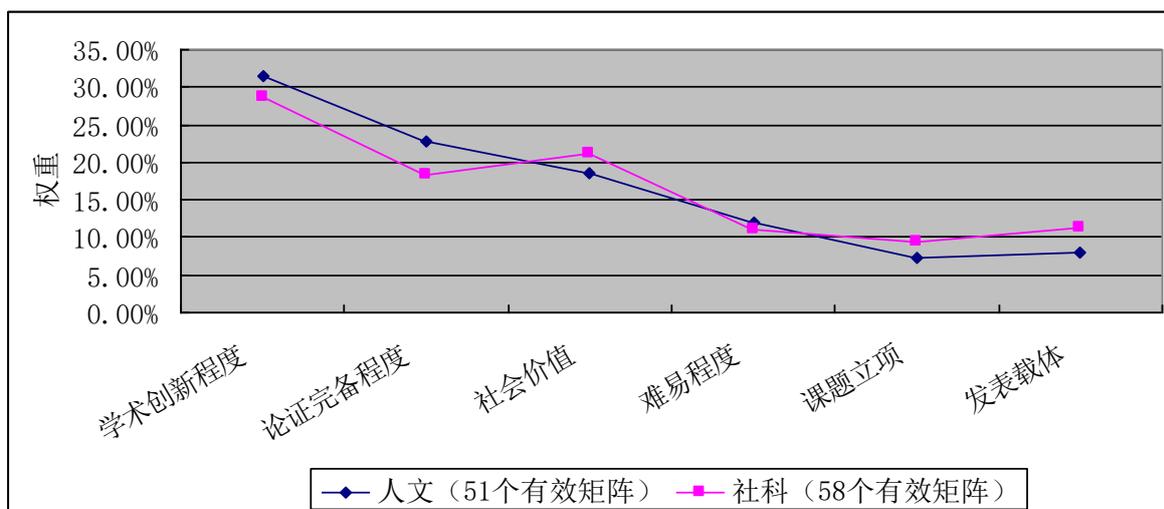


图 13 “人文学科论文”和“社会科学论文”指标权重分配对比

#### 4.2.4.3 “综述文章”与“研究论文”的指标权重分配存在明显差异

如图 14 所示,对选择方案二的受访者进行统计分析发现,“综述文章”和“研究论文”在指标权重分配上有如下差异:

- “综述文章”比“研究论文”更偏重“辅助指标”;

➤ “研究论文”的“学术创新程度”指标的权重明显高于“综述文章”。

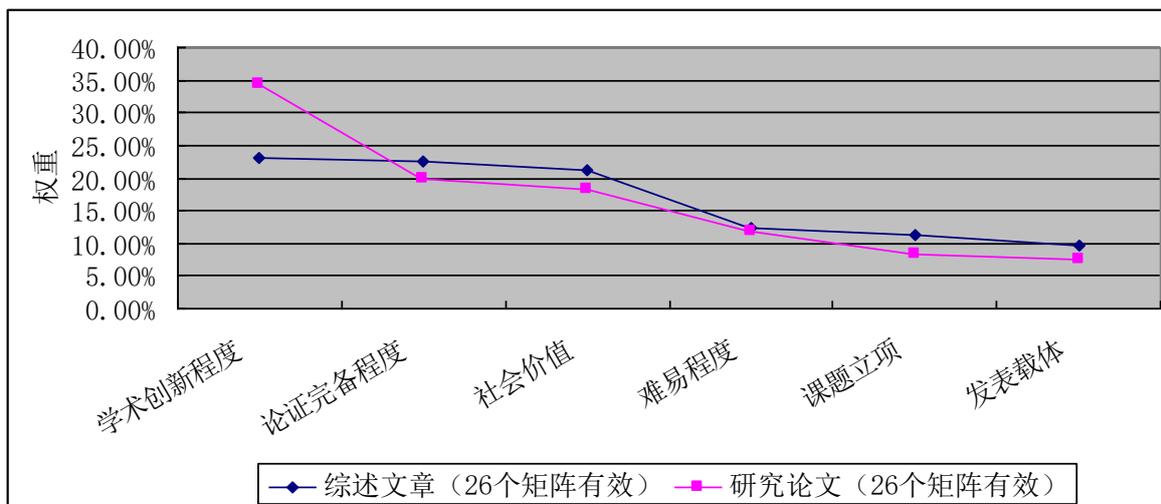


图 14 “综述文章”和“研究论文”指标权重分配对比

#### 4.2.4.4 “基础理论研究论文”与“应用实证研究论文”的指标权重分配基本一致

如图 15 所示，对 176 个有效矩阵中的“基础理论研究论文”和“应用实证研究论文”的矩阵进行分别统计分析，发现两类论文在指标权重分配上差异很小，可忽略不计。

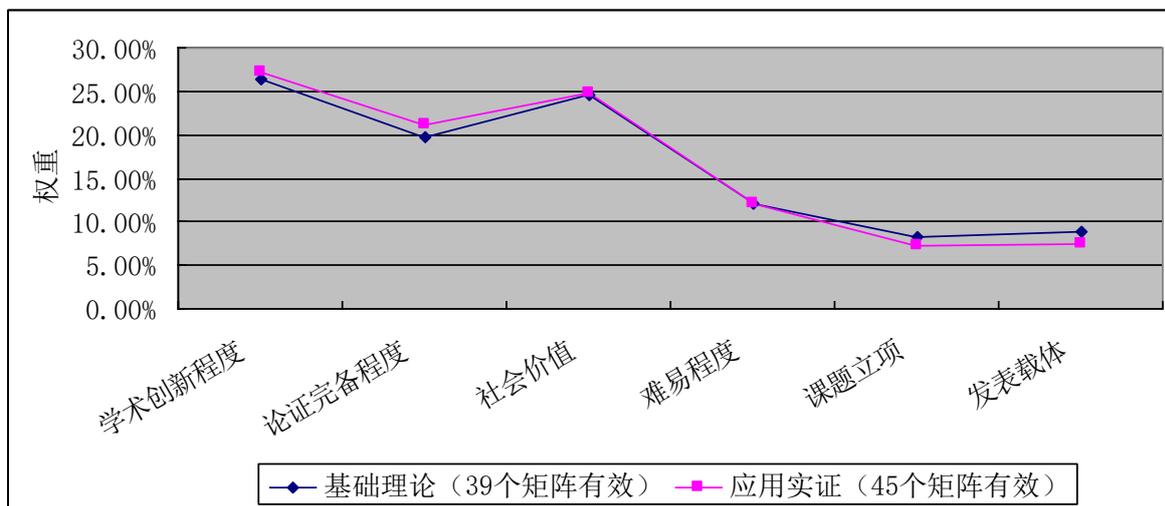


图 15 “基础理论研究论文”和“应用实证研究论文”指标权重分类对比

#### 4.2.4.5 不同学科背景的受访者所认同的指标权重分配存在一定的差异

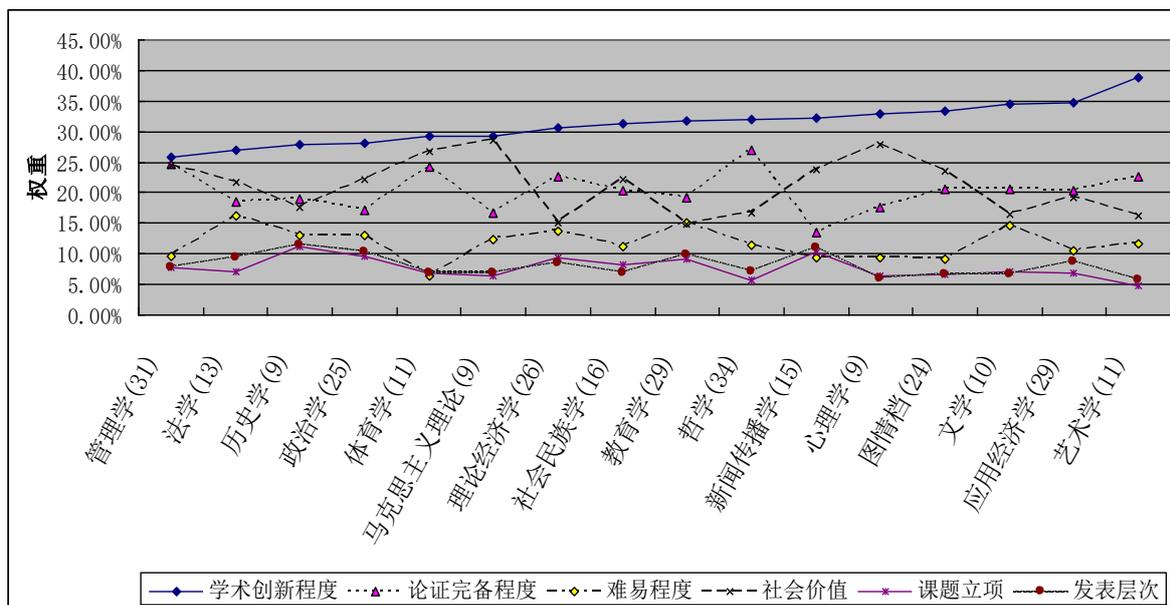


图 16 各学科指标权重分配对比

(注：括号内数字为“有效矩阵数量”)

如图 16 所示，各学科受访者普遍认为，“学术创新”、“论证完备”、“社会价值”是各学科论文质量评价体系中较为重要的指标。“学术创新程度”是衡量人文社会科学论文质量的首要指标，权重值在 26%-39%之间，远远高于其他指标的权重。其次是“论证完备程度”（13.38%-26.96%）和“社会价值”（14.91%-28.45%）的权重值相差无几。“难易程度”指标的权重排在第四位，权重值处于 6%-16%之间，前三个指标差异较大，略高于“发表载体”和“课题立项”的权重值。“课题立项”与“发表载体”权重取值范围各在 0.05-0.10 之间，但“发表载体”的权重在大多数学科中略高于“课题立项”的权重。单就各项指标权重的差异而言：

- “学术创新程度”的指标权重：虽然各学科受访者均认为该指标的权重应最高，但在具体权重值的设置，存在明显差异：其中艺术学学科背景的受访者所认为的“学术创新程度”指标权重值最高，为 38.89%，

管理学学科背景的受访者所认为的“学术创新程度”指标权重值最低，为 25.90%。

- “论证完备程度”的指标权重：各学科背景的受访者对“论证完备程度”的看法差异较大：其中，哲学学科背景的受访者分配该项指标的权重值最高，为 26.96%；管理学、体育学、理论经济学学科背景的受访者也赋予了该项指标较高的权重；然而心理学、政治学、马克思主义理论、新闻传播学学科背景的受访者，对该项指标权重赋予了较低的权重，其中新闻传播学的权重值最低，为 13.38%。
- “难易程度”的指标权重：在各学科中，法学学科背景的受访者最为看重“难易程度”对论文质量的影响，赋予该指标 16.33%的权重，其次是教育学和文学学科背景的受访者，也赋予了较高的权重；不过管理学、心理学、新闻传播学、图情档、体育学等专业学科背景的受访者赋予该项指标的权重，普遍低于 10%，其中体育学受访者赋予的权重最低，为 6.35%。
- “社会价值”的指标权重：社会科学各学科的受访者赋予了该项指标较高的权重，其中，马克思主义、心理学、体育学所赋予的权重最高，均高于 26.68%；而人文学科背景和理论性较强的学科，赋予该项指标的权重较低，比如历史学 17.56%、哲学 16.77%、文学 16.50%、艺术学 16.30%。其中较为特殊的是教育学，虽然作为社会科学，但赋予该指标的权重为 14.91%。
- “课题立项”的指标权重：历史学和新闻传播学学科背景的受访者，最为看重“课题立项”对衡量论文质量的作用，分别赋予 11.24%和 10.28%的权重；然而，哲学和艺术学学科背景的受访者，分别赋予

该项指标 5.62%和 4.14%的权重。

- “发表层次”的指标权重：该项指标的学科差异与“课题立项”指标类似，历史学和新闻传播学学科背景的受访者赋予了最高的权重，分别赋予 11.41%和 11.05%的权重；然而，心理学和艺术学学科背景的受访者，分别赋予该项指标 6.04%和 5.74%的权重。

#### 4.2.4.6 结论分析及最终权重分配方案

从总体调查结果上看，各项指标的重要性程度从高到低可分为四级，依次为：

- 学术创新程度
- 论证完备程度、社会价值
- 难易程度
- 发表载体、课题立项

此外，问卷调查显示，指标的权重设置在论述体裁和学科背景两个方面存在较大显著差异；按照研究性质划分的“基础理论研究论文”和“应用实证研究论文”的权重设置基本一致。因此，经过课题组反复讨论和通过对专家咨询求证，形成如下论文质量评估的指标归类和权重分配设置方案：

- (1) 将主要指标与辅助指标的权重比例按照综合统计分析结果，统一确定为“85%：15%”。
- (2) 为综述文章单独设置权重分配方案，见表 4。
- (3) 各学科论文设置不同的权重分配方案，见表 5。

## 附件：调查问卷

### 您的基本信息

姓名：\_\_\_\_\_

研究学科：\_\_\_\_\_

Email：\_\_\_\_\_

手机：\_\_\_\_\_

职业类型：\_\_\_\_\_ 1.教学研究 2.编辑 3.科研管理 4.其它\_\_\_\_\_

### 问卷一：论文分类方案的调查

【说明】对论文进行科学合理、易于操作的**分类并分别制定评价指标**，有助于增强评价结果的可比性。本问卷提供了四种论文分类方案，是否科学、可行，请您提出宝贵意见。

(一) 以下对论文的分类哪种更为合理？（若多选请排序）：\_\_\_\_\_

您的选择理由是：\_\_\_\_\_

方案 1 按学科大类划分：             人文学科论文             社会科学论文

方案 2 按论文的论述体裁划分：    研究论文                 文献综述（含述评）

方案 3 按科研活动的性质划分：    基础理论研究论文    应用实证研究论文

方案 4 同一指标，区别权重：构建一个基本的评价指标集合，根据不同学科（政治学、经济学、管理学、教育学、历史学、哲学、文学等）对指标赋予不同权重。

(二) 您的其它分类建议：\_\_\_\_\_

### 问卷二：论文评价指标的调查

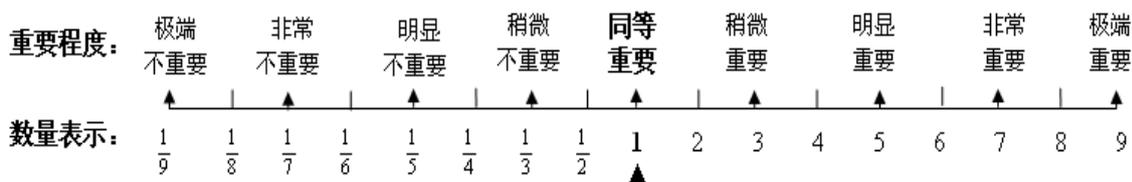
【说明】根据您在[问卷一]中选择的分类方案，判断“评价内容”对考量该类论文学术质量的必要程度：**1.很必要，2.有必要，3.必要程度不大，4.不必要**；填序号即可。未选分类方案不用填写。

评价指标	评价内容	方案 1		方案 2		方案 3		方案 4
		人文 学科	社会 科学	研究 论文	文献 综述	基础 理论	应用 实证	按您所在 学科填写
学术创新 程度	提出新的（或修正完善已有的）学说、理论、观点、问题、阐释等，对学术发展的促进程度							
	提出新的（或改进运用已有的）方法、视角等，对学术发展的促进程度							
	发现新的资料、史料、证据、数据等，对学术发展的促进程度							
	对已有成果做出新的概括、评析，对学术发展的促进程度							
论证完备 程度	理论前提的科学性							
	研究方法的有效性							
	论据的充分可靠性							
	逻辑的严密系统性							
	学术引证的规范性							
难易程度	论题的复杂程度（理论难点或实证研究难度）							
	资料搜集或处理的难度							
社会价值	对解决经济、政治、社会建设中问题的推动作用							
	对思想道德文化建设的促进作用							
辅助 指标	提出新的（或修正完善已有的）学说、理论、观点、问题、阐释等，对学术发展的促进程度							
	课题立项 论文来源的课题立项情况，如国家社科规划课题，省部级重点课题等							
	发表载体 论文发表载体的学术影响力，如 SCI、SSCI、A&HCI 来源刊、光明日报、人民日报，CSSCI 来源刊、中文核心期刊、中国人文社会科学核心期刊等							
[请补充]								

其它指标设置意见： \_\_\_\_\_

### 问卷三：论文评价指标权重的调查

**【说明】**：本问卷将用“层次分析法”调查评价指标之间的重要程度。请您将指标两两比较，按照下图中“重要程度”对应的“数量表示”赋值。如：指标 X 比指标 Y 非常重要，数量表示为“7”；指标 X 比指标 Y 非常不重要，数量表示为“ $\frac{1}{7}$ ”。



\* 打分示例如下表，纵向指标与横向指标相比：

- 若指标 A 比指标 B 明显重要，在 A 行和 B 列交叉位置填入“5”；
- 若指标 A 与指标 C 同等重要，在 A 行和 C 列交叉位置填入“1”；
- 若指标 A 与指标 D 重要性相比，处于稍微重要和明显重要之间，在 A 行和 D 列交叉位置填入“4”；
- 若指标 B 比指标 D 稍微不重要，在 B 行和 D 列交叉位置填入“1/3”。

↻	A	B	C	D
A		5	1	4
B			1/5	1/3
C				4
D				

请根据您的[问卷一]中选择的分类方案，填写下表。未选分类方案以及黑底斜线部分不用填写。

→若您选择了方案 1，请填写以下两个表

人文学科论文评价指标的比较

(★注意：若指标 A 比 B 重要、B 比 C 重要，则 A 比 C 重要。)

↻	学术创新	论证完备	难易程度	社会价值	课题立项	发表载体
学术创新						
论证完备						
难易程度						
社会价值						
课题立项						
发表载体						

社会科学论文评价指标的比较

(★注意：若指标 A 比 B 重要、B 比 C 重要，则 A 比 C 重要。)

↻	学术创新	论证完备	难易程度	社会价值	课题立项	发表载体
学术创新						
论证完备						
难易程度						
社会价值						
课题立项						
发表载体						

→若您选择了方案 2，请填写以下两个表

研究论文评价指标的比较

(★注意：若指标 A 比 B 重要、B 比 C 重要，则 A 比 C 重要。)

↻	学术创新	论证完备	难易程度	社会价值	课题立项	发表载体
学术创新						
论证完备						
难易程度						
社会价值						
课题立项						
发表载体						

文献综述（含述评）评价指标的比较

(★注意：若指标 A 比 B 重要、B 比 C 重要，则 A 比 C 重要。)

↻	学术创新	论证完备	难易程度	社会价值	课题立项	发表载体
学术创新						
论证完备						
难易程度						
社会价值						
课题立项						
发表载体						

→若您选择了**方案 3**，请填写以下两个表

基础理论研究论文评价指标的比较

(★注意：若指标 A 比 B 重要、B 比 C 重要，则 A 比 C 重要。)

↻	学术创新	论证完备	难易程度	社会价值	课题立项	发表载体
学术创新						
论证完备						
难易程度						
社会价值						
课题立项						
发表载体						

应用实证研究论文评价指标的比较

(★注意：若指标 A 比 B 重要、B 比 C 重要，则 A 比 C 重要。)

↻	学术创新	论证完备	难易程度	社会价值	课题立项	发表载体
学术创新						
论证完备						
难易程度						
社会价值						
课题立项						
发表载体						

→若您选择了**方案 4**，请您根据所在学科特点填写下表

(★注意：若指标 A 比 B 重要、B 比 C 重要，则 A 比 C 重要。)

↻	学术创新	论证完备	难易程度	社会价值	课题立项	发表载体
学术创新						
论证完备						
难易程度						
社会价值						

课题立项						
发表载体						

其它权重设置意见： \_\_\_\_\_

## 4.3

# 人文社会科学论文质量评估指标体系 专家论证报告

中国人民大学人文社会科学学术成果评价研究中心

2010年12月10日

### 4.3.1 专家论证的研究概述

2010年10月22日-2010年12月3日之间,课题组通过电子邮件、电话和面访等方式,邀请了约30位<sup>89</sup>学术评价领域的权威理论专家学者和资深实践先行者(专家名单见本报告附件一),就设计的《人文社会科学论文质量评估指标体系》(以下简称《指标体系》)进行了咨询和论证(《专家论证评审书》见本报告附件二和《咨询的主要问题》见本报告附件三)。各位专家对《指标体系》的设计理念与方法、分类评估方案、基本指标设置、权重分配方案、组织实施方案等,进行了全面系统深入的分析,并有针对性提出了具有重要参考借鉴价值的论证意见和改进建议。

课题组通过汇总分析专家的论证意见形成本论证报告,以作为《指标体系》修改完善和最终定稿的重要依据。

### 4.3.2 指标体系的整体论证

多位专家指出,以“同行评议为主,文献计量为补充”的操作思路切合当前评价的主流思想,体现了人文社会科学研究的特点(8人);《指标体系》以论文为直接评估对象,有助于宏观研究和评价分析,具有创见性和实用价值(3人);《指标体系》的构建,有利于满足人文社会科学论文评估的迫切需求,体系结构完整、思路缜密、观点清晰,论证过程较为科学合理、有说服力(10人)。

部分专家(7人)指出,总体上《指标体系》具备一定的可操作性,但“同行评议”部分的实施方案有待改进;由于工作量非常大,需要考虑必要的筛选手段和实施成本。

【课题组分析】《指标体系》的创新性、理论基础和逻辑框架通过了专家的检验;《指标体系》的实施成本和可操作性有待进一步改进。

### 4.3.3 分类评估方案的论证

<sup>89</sup> 其中18位正式反馈,其他专家提出了口头或电子邮件等非正式形式的建议和意见。

#### 4.3.3.1 关于整体分类评估方案

多数专家（9人）指出，《指标体系》的分类评估方案较为合理，尤其是对综述文章单独设置权重分配方案的观点可取。有的专家（4人）指出，“统一设置指标，按一级学科分别设置权重、按二级学科进行同类比较”的思路，比设置多套指标更为科学合理和可行性。

部分专家认为《指标体系》的分类方案，可在如下几方面精益求精：

- 在坚持“学科分类评价”的大原则下，应根据评估具体需要，灵活掌握分类；尤其是“二级学科”的分类，可根据学科整体论文数量大小进行细化或合并（2人）；若有必要，也可针对二级学科单独设置权重（1人）。
- 综述文章在指标上也应有差别于其他一般研究论文，比如去掉“学术创新程度”（2人）。
- 针对偏定量研究的论文，应单独设一类进行评估（1人）。
- 指标或权重设置应考虑论文在“理论研究”和“应用研究”方面的差异性（2人）。
- 针对译文、学术动态、领导讲话、读者来信、卷首语等特殊类别的论文，应单独设一类，进行特别处理（2人）。
- 把施引论文所在期刊的分类，作为论文分类依据（1人）。

【课题组分析】鉴于专家对当前分类方案的基本认同，《指标体系》将保留原有的分类框架，并做如下改进：

- 1) 根据二级学科论文数量和评估实际需要，形成论文“二级学科分类表”，作为同类比较的依据；
- 2) 对于专家提出的部分论文需要单独分类、单独设置指标或权重的建议，将通过实证检验，确定是否推行，若这些论文的差异性不明显，或论文总量太小，为了操作便利性，将不再进行单独分类。

#### 4.3.3.2 关于跨学科论文分类评估方案

在整体分类评估方案中，针对如何处理跨学科论文的分类，专家的观点分为

如下三种：

- **单独设置跨学科类**（共 3 人）：1) 按大类分为人文跨学科、社科跨学科和艺术跨学科（1 人）；2) 按照跨学科主题进行分类（2 人）。
- **将跨学科论文划分到各学科进行分别评估**（共 3 人）：1) 以作者所在的研究学科为准（1 人）；2) 根据论文内容涉及学科比例最高的，确定论文的学科（1 人）；3) 按照论文内容所解决问题的所属学科，确定论文的学科（1 人）。
- **综合分类（共 2 人）**：以某一学科的讨论为主的，可以直接归到该学科评估；确实分不清的，可归入单设的“跨学科”类（2 人）。

若设置单独的跨学科类，该类论文的评估指标权重可采用两种方法确定：1) 由相应的跨学科专家论证提出（1 人）；2) 根据文章所涉及的学科类别及相应权重按比例计算得出（1 人）。

【课题组分析】对于跨学科论文的分类，最终的方案如下：

- 1) 围绕某一主要学科的理论、方法或问题开展研究，但涉及其他学科内容。此类论文在该主要学科中进行评估，形成唯一评估结果。
- 2) 研究同时涉及一个以上学科的理论、方法或问题，难以归属到具体学科。此类论文同时在不同学科中进行独立评估，形成多个“学科评估结果”和一个“综合评估结果”；后者以各学科评估中的最高得分为准。

## 4.3.4 基本指标设置的论证

### 4.3.4.1 关于同行评议指标

《指标体系》提出“从论文内容质量内涵的 4 个维度出发”设计同行评议指标，大部分专家对这一理念比较认同，但有 4 位专家对这一理念提出质疑：

- “正确”这个维度缺乏明确标准（1 人）；
- “新颖”这个维度可改为“创新”或“前沿”（2 人）；
- 这 4 个维度是学术论文基本的标准，不宜作为理论基础（1 人）。

【课题组分析】关于设计理念，可反映在同行评议指标的含义解释上，《指标体系》不再把这部分作为理论基础进行重点阐述。

对于同行评议指标中的4个主要指标，专家普遍比较认同，针对其中的3个指标的名称和含义提出了微调建议：

- **“论证完备程度”**：建议改为“研究规范程度”（1人）；
- **“难易程度”**：评估内容包含的“资料收集或处理难度”，没有必要或不准确（3人）；指标名称建议改成“工作量”（1人）；建议取消“难易程度”指标，因为“学术创新程度”包含了这一内容，若保留可能把“简单问题复杂化”（1人）。
- **“社会价值”**：对指标内涵的界定略显广泛（1人）。

【课题组分析】鉴于专家对同行评议主要指标的设置普遍较为认同，且这4个指标已经过广泛的文献调研和问卷调查，因此暂时不做修改。

对于同行评议指标中的2个辅助指标，专家提出了如下改进建议：

- **“课题立项”**：应对横向和商业项目的归类标准加以说明（1人）；应对国际项目的资助机构等进行更为细致的规定（2人）；省部级以下项目的论文质量未必高，省部级以下的项目不应归类（1人）。
- **“发表载体”**：A级载体中还应考虑《新华文摘》和《中国社会科学》，建议发表载体分为三级：国外收录+中国社会科学+人民日报、光明日报，核心期刊，非核心期刊（1人）。
- “课题立项”和“发表载体”不属于同行评议指标，建议与文献计量指标放在一起（1人）；
- 建议删除“课题立项”和“发表载体”指标，因为仍体现了“以刊评文”，尤其是立项基金标注的随意性很大（2人）；

【课题组分析】《指标体系》坚持以“直接指标为主，间接指标为辅，文献计量指标为补充”，之所以将这两个指标纳入，主要是考虑把论文形成和发表过程中的“同行评议结果”作为参考；由于这两个辅助指标经过充分的分析论证，且有13位专家建议保留，因此《指标体系》仍保留这两个指标；同时，还将参照上述改进建议，对这两个指标的评分标准进行修正和细化（见2.2.1）。

关于“难易程度”指标用“系数”还是“权重”表示，专家的争议较大：

一部分专家认为（5人），用系数更为合理和准确；系数的调节范围可根据样本大小、专家的平均值、运算需要等条件确定。

另一部分专家认为（5人），用权重更为稳妥；因为《指标体系》更看重评估结果的整体（而不是看重难度），“难易程度”指标发挥的作用有限，若用系数，容易造成放大效应，较难控制；另外，若用系数表述，也要对计算依据、取值范围等进行细致论证，对数学符号的定义也应更为严密。

【课题组分析】对此问题专家争议较大，但相比来讲，用“权重”表示更为“稳妥”，用“系数”表示更为“复杂”。《指标体系》暂定为用“权重”表示，在实证中需再对“系数”表示的可行性进行进一步试验。

#### 4.3.4.2 关于文献计量指标

专家对《指标体系》中的文献计量指标的设置，提出了较为细致的意见和建议。包括如下三类：

##### (1) 文献计量指标的总体设计思路。包括：

- 多数专家认为（5人），改造评估期刊指标用于评估论文不合适；少数专家认为（1人）也可进行改造，但要通过深入研究和专家论证。
- “文献计量指标”作为“补充”，不宜过细，下载和被引两个指标足以（1人）。
- 改为“科学（文献）计量学评估指标体系”（原称“引证计量指标体系”）（1人）。
- 多位专家指出，对每个文献计量指标的计算方式、限定要素等，应进行更为详细和明确地界定，例如数据获取方式、时间区间、交叉要素等。
- 文献计量指标的“年均”概念无意义，应计算3年引文数据的总量（2人）。

##### (2) 具体文献计量指标的修改。包括：

- **Web 下载率**：建议名称修改成“年均 Web 下载频次”（1人）；为避免人

为增加，应注意技术上如何辨别同一人多次下载（1人）；应注意统计方面的科学性（1人）。

- **年均被他引频次**：对“他人引证”应明确“仅限第一作者”还是“所有合著作者”（1人）；引用文献的质量也应予以考虑，如施引主体的学科特征、等级权重，引用目的等（2人）。
- **论文被引广度**：不按学科计算，按机构计算，因为小学科中引证广度的区分度很小（1人）；可以增加学科、期刊的广度（2人）；“论文被引广度”似乎与论文的质量关系不大，建议取消（1人）。
- **论文 h 指数**：理论依据不充分，大部分论文区分度较小，建议暂缓纳入（2人）；很可能区分度较小，可先做一些实证研究，再作定论（2人）；建议指标名称应改为“影响力指数”（1人）；建议指标名称改为“权威作者或高影响因子期刊引用指数”（1人）；该指标对年轻学者不公平（1人）。

### **(3) 建议增加文献计量指标。包括：**

- 即年被引指标、参考文献数和作者有关的一些指标、单篇论文的学科贡献度（可以用 Pagerank 等新指标）（1人）；
- 基金资助指标（1人）；
- 自引也不是都不好，建议增加总被引频次指标（1人）。

**【课题组分析】**综合上述意见和建议，课题组决定：

- 1) 暂时保留现有的 3 个计量指标，Web 下载频次、在实证中通过数据最终确定指标是否保留；
- 2) 专家提出增加的指标，普遍处于理论研发阶段，出于数据获取、实施操作等方面的考虑，暂时不纳入体系；
- 3) 由于论文引证数据太少，取消指标的“年均”概念，代之以“总量”概念；为了平衡不同论文上网时间不同导致的不公平，“Web 下载频次”以“月均”的概念计算；
- 4) 采纳专家的建议，修改部分指标的名称，细化说明指标的限定条件。

#### 4.3.4.3 关于两类指标的关系

关于“同行评议”和“文献计量”两类指标的关系，专家的意见可分为如下两类：

**(1) 两种评价方法是相互补充的关系，相加依据不足（共5人）：**不必追求“相对一致性”，两者有时需要“一票否决”；可分别统计三个榜：同行评议认为的优秀文章，文献计量计算出的优秀文章，二者都上榜的文章（4人）；上五年的计量结果，可作为下一轮的同行评议专家的参考，不必把二者绝然分开（1人）；不宜称作“即时和延时评价相结合”，而应看作是“定性与定量评价相统一”（1人）

**(2) 两种评价结果在一定条件下可以相加（共5人）：**统计时间区间一致的情况下，可以相加（3人）；量纲归一化之后再合成（1人）；在明确评估目的后，才能最终确定二者关系（1人）；可对两类指标的评估结果，做进一步的实证研究，发现其内在关系，比如两者在哪一时间段内相关系数最高（1人）。

【课题组分析】综合参考上述两类意见，通过实证确定两类指标的最终关系。

#### 4.3.5 权重分配方案的论证

关于“同行评议”和“文献计量”两类指标权重的分配问题。认为两者可相加的专家（5人）认为，既然“同行评议为主，文献计量为补充”，则应前者占主体，并与后者适当拉开差距。若两者评估结果相似度不高，有必要结合（1人）；权重应经过实证检验和细致测算（2人）。具体权重比例分为如下三类：

- 同行评议 70%，文献计量 30%（5人）；
- 同行评议 40%，文献计量 60%（3人）；
- 同行评议 20%，文献计量 80%（1人，参考英国 TIMES 大学排名）。

关于同行评议指标的内部权重分配，只有1位专家提出：综述文章的“学术创新程度”权重可提高一些，“论证完备程度”权重可降低一些。

关于文献计量指标的内部权重分配，专家提出的意见主要包括：

- 年均被他引频次，是最重要的指标，权重应最高；其他指标权重差不多（1人）。
- 被《新华文摘》等转载的指标，权重还可下降些（1人）。
- Web 下载指标的不确定因素更多，权重不应太高（1人）。

**【课题组分析】**

- 1) 针对当年数据的评估结果，两类指标的权重分步为：同行评议 85%，文献计量 15%；针对 3 年数据的评估结果中，两大类指标的权重分配：同行评议为 70%；文献计量为 30%；通过实证确定最终分配结果。
- 2) 同行评议内部指标的各学科权重做微调，重点控制“主要指标”和“辅助指标”的比例。
- 3) 文献计量内部指标权重暂时保持不变，待进一步实证研究。

### 4.3.6 体系实施方案的论证

#### 4.3.6.1 关于评审时间

关于评审时间，专家的意见主要包括：

- 文献计量指标 5 年足以（1人）。
- 对文献计量指标 5 年，进必要说明和论证（3人）。
- 考虑“同行评议”和“文献计量”两类评估结果的结合，5 年的计算周期太长，建议改为 3 年（2人）。
- 即时评价和延迟评价的概念需要更细致地阐述，文摘和下载率也有即时性（1人）。
- 在即时评估阶段，可加入一些当年可获取的文献计量指标，比如下载频次、转载数量等（2人）。
- 请专家评估时，适当考虑时间影响（1人）。
- 必须在同一时间窗内进行比较和评价（3人）。

**【课题组分析】**

- 1) 将文献计量的统计时间区间改为“3 年”，并对选择理由进行说明。

- 2) 去掉“即时评估”的概念，将统计周期改为1年和3年两个时间点。
- 3) 其他意见将通过实证研究确定是否进行修改。

#### 4.3.6.2 关于数据来源

关于数据来源，专家的建议包括：

- 建议与数据量大的数据提供商合作，数据量太小会影响结果准确性（1人）。
- 尽量减少对外部数据源的依赖（1人）。
- 报告应将评估范围界定为“报刊”而非“期刊”，以便包含报纸载体的论文；对评估对象的数量进行必要地估算（2人）。

【课题组分析】实施过程中将充分参考上述意见。

#### 4.3.6.3 关于专家主观性的控制

多位专家（8人）对同行评议中“专家主观性的控制”提出了解决方案：

- 应对论文评价结果的用途和时效性进行说明（1人）。
- 采取措施，督促专家在一定的时限内完成评审；（2人）。
- 采用单向盲审（1人）。
- 建议在某些学科先做实验（2人）。
- 指标的评分表，应提供更为详细的评判标准（1人）；
- 建立专家库，建立同行评估专家团队的组成规则和操作方式，确保选择的专家诚信、可靠、责任心强并有充足的评审时间（3人）。

【课题组分析】参考上述意见，《指标体系》的实施将采用以下措施：

- 1) 细化评分表内容，使之更为可操作；
- 2) 建立专家库和相应的评审工作制度；
- 3) 利用信息技术分析评议行为，在合成评分结果时，采用数学合成方法以最大可能性地消除专家主观性所带来的负面影响。

#### 4.3.6.4 关于大规模同行评议可行性

共有 7 位专家针对大规模同行评议的可行性问题，提出了解决建议，包括：

- 按照优秀论文的离散分布归类，通过“粗评”去掉“长尾”（质量一般的论文），只对精华论文进行细评（6 人）；分成杰出和可能杰出、优秀和可能优秀及普通三层，第一层是千分之一~千分之二，第二层是百分之一~百分之二；只评价前两类（1 人）。
- 估算各学科每年可能产生的大致论文数量，聘请与之相对应的专家数量（2 人）。
- 采用网络评价，依托计算机软件进行全部计算（1 人）。
- 由期刊杂志自荐或他荐来筛选论文（3 人）；推荐论文可进行公开，约束不合理的推荐（1 人）。
- 每年只评估部分学科，轮流评估（1 人）。
- 重点评价各编辑部放在封面上介绍的文章（1 人）。
- 明确评价目的（如所有已发表论文的选优，还是及时推荐相对有特点的论文），先确定好刊（以刊评文），设置文献计量门槛（1 人）。
- 引入查新机制（1 人）。

【课题组分析】课题组主要参考了上述前三条建议：

- 1) 对同行评议论文的数量进行估计；
- 2) 设置了多轮筛选方案，重点对精华论文进行评估；
- 3) 研发软件系统完成所有统计计算工作，提高工作效率。

**附件一：评审专家名单（按姓氏拼音排序）：**

- **安小米：**中国人民大学信息资源管理学院，教授、博士生导师
- **蔡蓉华：**北京大学图书馆文献计量学研究室原主任，研究馆员
- **樊秀娣：**同济大学办学质量评估院，室主任/副研究员
- **何峻：**北京大学图书馆文献计量学研究室主任，副研究馆员
- **纪光华：**北京师范大学数学科学学院，副教授，硕士生导师
- **姜晓辉：**中国社会科学院文献信息中心文献计量学重点研究室，主任、研究员
- **刘益东：**中国科学院自然科学史研究所，中国科学院学部学科发展战略研究中心主任、研究员、博士生导师
- **庞景安：**中国科学技术信息研究所，研究员、硕士生导师
- **苏新宁：**南京大学信息管理系，教授、博士生导师
- **武夷山：**中国科学技术信息研究所，总工程师 / 研究员、博士生导师
- **杨立英：**中国科学院文献情报中心，副研究员
- **叶继元：**南京大学中国人文社会科学评价国家创新基地副主任，南京大学信息管理系教授、博士生导师
- **俞立平：**宁波大学商学院，教授
- **袁军鹏：**中国科学技术信息研究所，研究员、硕士生导师
- **赵丹群：**北京大学信息管理系，副教授、硕士生导师
- **王星：**中国人民大学统计学院，副教授、硕士生导师
- **周萍：**中国科学技术信息研究所，研究员、硕士生导师
- **周庆山：**北京大学信息管理系，系副主任，教授、博士生导师

**附件二：《人文社会科学论文质量评估指标体系》评阅书**

**对该评估体系的整体评价意见**（请按以下两个方面填写）：

- (1) 该评估体系值得理论肯定和实践采纳之处。
- (2) 该评估体系尚存在的问题及改进建议，特别是推进其应用操作的指导意见。

签字（盖章）：

年 月 日

以下请专家填写：

工作单位	职务/职称	联系电话

（可加附页）

### 附件三：需要咨询专家重点解答的问题

#### 在阅读本《体系》时，有以下六个问题需要得到您的指导：

(1) 论文的分类评价方案（统一指标，综述文章单独设置权重，其他研究论文按一级学科设置不同的权重）是否合理？您认为跨学科的论文，应如何分类比较好？

(2) 当前，针对论文本身评估的引证计量指标极少，基本都是评估期刊的指标；且论文的引证数据量偏小，此时如果改造评估期刊指标用于评估论文，是否有其可行性和科学性？除报告里列出的指标外，是否还有其它可用于单篇论文的引证计量指标？

(3) 难易程度指标的计算结果是用权重还是用系数来表示呢？如果用难易系数表示，难易系数怎么来确定呢？我们在报告第 22 页中给出了难易程度的计算公式，您觉得这个系数在多大范围内调节比较合适？

(4) 同行评议是定性的即时评价（一年）方法，引证计量是定量的延时评价（五年）方法，为了保证两者的评价结果具有相对一致性和校验佐证性，两者在设计各自的评价指标时，需要注意什么问题？

(5) 引证计量评价结果与同行评议评价结果的合成目前采用的是三七开再相加，您觉得这两部分的评价结果可以直接相加吗？如果能，权重比例各为多少比较合适？

(6) 评估的论文数量非常大（每年约 50 万篇），您觉得实施方案需要注意哪些问题，以便提高工作效率？

## 4.4

# 人文社会科学论文质量评估系统（2.0）

## 使用指南

中国人民大学人文社会科学学术成果评价研究中心

2011年12月10日

#### 4.4.1 登陆

用户直接访问人文社科论文质量评估系统，书报资料中心的编辑可在中心局域网上访问网页 <http://218.240.133.151:8080/rssi2/>，弹出登录界面（如图 17 所示）。输入用户名和密码，单击“登录”按钮，即可进入评估系统 2.0。



图 17 系统登录界面

#### 4.4.2 菜单管理

登录后，左侧菜单栏共有五个功能模块：基本信息管理、评审管理、统计分析、信息查询、系统管理。

- “基本信息管理”功能：为用户及评价人员提供论文评价的后台数据管理功能；后台数据包括课题立项、报刊管理、权重管理、二级学科、地区管理、专题信息的分级管理；本模块提供这些数据的查看、修改、增加、删除等功能。
- “评审管理”功能：提供论文的评审功能，包括论文的分类管理、评分管理、评分分布、评分明细等。
- “统计分析”功能：为评价分析人员提供汇总分析每篇论文得分的功能，在此模块，评价分析人员可通过设置不同的统计分析条件，统计得出期刊、作者单位、作者个人、地区等论文的总体得分情况，并可自动生成统计报表。
- “信息查询”功能：提供评审进度查询功能，包括对各评委&专题论文数查询、评委论文数查询、专题论文数查询、每日评论论文数的查询。
- “系统管理”功能：提供用户管理及密码修改功能。用户可通过此模块修改密码，密码可使用数字和字母，字母应区分大小写，密码长度没有限制。如果离开页面半个小时以上，系统会返回登录界面，用户需要重新登录。

### 4.4.3 基本信息管理

点击左侧的“基本信息管理”，下方显示六个可点击的按钮，依次是“课题立项”、“报刊管理”、“权重管理”、“二级学科”、“地区管理”、“专题信息”。此页面仅为用户提供查看浏览的功能，不提供修改、增加、删除等权限；评价研究人员具有该模块中修改、增加、删除等功能，负责对本模块进行更新和维护。尤其是“课题立项”、“权重管理”以及“二级学科”的信息随着评价体系、学科的变化要及时做出调整；其他信息随着核心期刊体系和期刊名称、版别的变更，在每年年初做相应维护。

#### 4.4.3.1 “课题立项”基本信息管理

##### (1) 查看“课题立项”的内容

单击“课题立项”，如图 18 所示，页面显示课题名称与相应分数的列表。比如，一篇论文属于国家级课题资助，则“课题立项”指标的得分为 21 分。



图 18 基本信息管理-课题立项界

##### (2) 修改“课题立项”的内容

在要修改的课题立项所在行点击“编辑”，显示“修改课题立项”对话框，如图 19 所示，对话框显示修改前的内容，可直接在方框中输入修改后的内容，然后

点击提交，页面提示修改成功。

**修改课题立项**

* 代码:	01
* 名称:	国家级
* 分数:	21

图 19 课题立项管理 - 修改课题立项界面

### (3) 删除“课题立项”的内容

点击要删除的课题立项中某一课题级别所在行，再点击删除按钮。页面提示“是否确定删除”，单击“确定”则删除成功，单击“取消”表示删除失败。

### (4) 增加“课题立项”的内容

直接单击  按钮，页面显示“添加课题立项”界面，依次输入“代码”、“名称”、“分数”所在列的值，此处“代码”所在列只能输入两位数字，作统计用，无任何语义；输入完成之后单击提交按钮即可。

## 4.4.3.2 “报刊管理”基本信息管理

### (1) 查询“报刊管理”的内容

单击“报刊管理”，页面显示“报刊管理”界面，如图 20 所示，用户可根据检索需求选择相应的下拉菜单，点击按钮，即可显示出相应信息。



图 20 基本信息管理 - 报刊管理界面

例如，某用户需要查询“载体级别”为核心，“报刊品种”为期刊，“院校类型”为 211 院校的报刊信息，首先在“载体级别”选项的下拉菜单中选择“核心”，“报刊品种”选项菜单中选择“期刊”，“院校类型”菜单中选择“211 院校”，点击右边的“检索”按钮，即可显示出相应的期刊。如图 21 所示

基本信息管理 > 报刊管理

载体级别: 核心 | 报刊品种: 期刊 | 院校类型: 211院校 | 特色院校: 全部 | 报刊名称:

检索 添加 导出

报刊名称	版别	国际刊号	国内刊号	刊期	页码	发文数	出版地	载体级别	报刊内容	报刊品种	院校类型	特色院校	操作
《阿拉伯世界研究》		ISSN1673-5161	CN31-1973/C	双月	80	59	上海	核心	综合性	期刊	211院校	语言院校	编辑 删除
《安徽大学学报》	哲学社会科学版	ISSN1001-5019	CN34-1040/C	双月	156	125	安徽	核心	综合性	期刊	211院校	综合院校	编辑 删除
《北京交通大学学报》	社会科学版	ISSN1672-8106	CN11-5224/C	季刊	136	99	北京	核心	综合性	期刊	211院校	理工农医林院校	编辑 删除
《北京体育大学学报》		ISSN1007-3612	CN11-3785/G8	月刊	144	461	北京	核心	非综合性	期刊	211院校	体育院校	编辑 删除
《比较法研究》		ISSN1004-8561	CN11-3171/D	双月	160	85	北京	核心	非综合性	期刊	211院校	政法院校	编辑 删除
《财经科学》		ISSN1000-8306	CN51-1104/F	月刊	124	180	四川	核心	非综合性	期刊	211院校	财经院校	编辑 删除
《财经研究》		ISSN1001-9952	CN31-1012/F	月刊	144	155	上海	核心	非综合性	期刊	211院校	财经院校	编辑 删除
《东北师大学报》	哲学社会科学版	ISSN1001-6201	CN22-1062/C	双月	236	366	吉林	核心	综合性	期刊	211院校	师范院校	编辑 删除
《东疆学刊》		ISSN1002-2007	CN22-5016/C	季刊	112	0	吉林	核心	综合性	期刊	211院校	综合院校	编辑 删除
《东南亚研究》		ISSN1008-6099	CN44-1124/D	双月	96	0	广东	核心	综合性	期刊	211院校	综合院校	编辑 删除
《法商研究》		ISSN1672-0393	CN42-1664/D	双月	160	121	湖北	核心	非综合性	期刊	211院校	财经院校	编辑 删除
《高教发展与评估》		ISSN1672-8742	CN42-1731/G4	双月	124	110	湖北	核心	非综合性	期刊	211院校	理工农医林院校	编辑 删除

共[71]条记录 共[6]页 当前第[1]页 << 上一页 PRE 下一页 NEXT >>

图 21 基本信息管理 - 报刊管理界面

## (2) 修改“报刊管理”的内容

如图 21 显示属“211 院校核心期刊”的列表，鼠标点击《东北师大学报》，该选中行背景呈现绿色，点击“编辑”，显示“修改期刊”对话框，如图 22 所示，在方框中填入相应信息，点击提交，即可修改。

* 报刊名称：	《东北师大学报》
载体级别：	核心
版别：	哲学社会科学版
国际刊号：	ISSN1001-6201
国内刊号：	CN22-1062/C
刊期：	双月
页码：	236
发文数：	366
报刊内容：	综合性
报刊品种：	期刊
院校类型：	211院校
特色院校：	师范院校
出版地：	吉林

提交 重填 返回

图 22 报刊管理 - 修改期刊界面

(3) 增加“报刊管理”的内容



在“报刊管理”界面下，点击右上方按钮，弹出“添加报刊信息”对话框，如图 23 所示，单击方框可输入报刊名称、版别、发文数等信息，单击提交即可完成添加报刊信息。

* 报刊名称：	
载体级别：	核心
版别：	
国际刊号：	
国内刊号：	
刊期：	年刊
页码：	
发文数：	
报刊内容：	综合性
报刊品种：	期刊
院校类型：	非院校
特色院校：	师范院校
出版地：	安徽

提交 重填 返回

图 23 报刊管理 - 添加期刊信息界面

(4) 删除“报刊管理”的内容

单击要删除的报刊所在行，再单击“删除”。页面提示“是否确定删除”，单击“是”则删除成功，单击“否”表示删除失败。

## (5) 导出“报刊管理”的内容

如图 21 所示，单击页面右上角的“导出”按钮，即可导出 Excel 格式的报刊管理报表。

### 4.4.3.3 “权重管理”基本信息管理

在“基本信息管理”界面下，单击“指标权重”，显示图 24。其中，第一列为一级学科代码<sup>90</sup>，第二列为一级学科名称，第 3 列 - 第 8 列依次为在各综述文章或一级学科论文里，六个评价指标所占的权重。注意，前四个指标是主要指标，权重之和为 100%，占整个指标体系权重的 85%；后两个指标是辅助指标，权重之和也是 100%，占整个指标体系权重的 15%。如图所示，点击第一行“0000 综述文章……”，该行背景显示为绿色，通过“编辑”、“删除”、“添加”按钮，可以进行相应的修改，该操作与 4.4.3.1 部分“课题立项”信息管理部分一致，不再赘述。

基本信息管理 > 一级学科及权重								
代码	名称	学术创新	论证完备	社会价值	难易程度	课题立项	发表载体	操作
0000	综述文章	0.29	0.29	0.27	0.15	0.54	0.46	编辑
0101	哲学	0.38	0.28	0.18	0.16	0.33	0.67	编辑 删除
0201	理论经济学	0.36	0.27	0.21	0.16	0.53	0.47	编辑 删除
0202	应用经济学	0.41	0.24	0.22	0.13	0.4	0.6	编辑 删除
0301	法学	0.32	0.24	0.26	0.18	0.4	0.6	编辑 删除
0302	政治学	0.33	0.24	0.28	0.15	0.47	0.53	编辑 删除
0303	社会学	0.36	0.25	0.23	0.16	0.53	0.47	编辑 删除
0304	民族学	0.36	0.25	0.23	0.16	0.53	0.47	编辑 删除
0305	马克思主义理论	0.34	0.21	0.31	0.14	0.4	0.6	编辑 删除
0401	教育学	0.37	0.24	0.21	0.18	0.5	0.5	编辑 删除
0402	心理学	0.39	0.21	0.29	0.11	0.4	0.6	编辑 删除
0403	体育学	0.34	0.25	0.24	0.17	0.4	0.6	编辑 删除

共[31]条记录 共[3]页 当前第[1]页

图 24 基本信息管理 - 权重管理界面

### 4.4.3.4 “二级学科”基本信息管理

<sup>90</sup> 与教育部《授予博士、硕士学位和培养研究生的学科、专业目录》（1997 年颁布，人文社会科学部分）的学科代码完全一致，采用四位阿拉伯数字表示。此外为了保持代码格式一致，增加了“综述文章”的代码为“0000”。

在“基本信息管理”界面下，单击“二级学科”按钮，显示图 25。其中，第一列为二级学科代码，第二列为二级学科名称<sup>91</sup>。

二级学科的编辑、增加、删除功能，与 4.4.3.1 部分“课题立项”信息管理部



分一致，不再赘述。

图 25 基本信息管理 - 二级学科管理界面

#### 4.4.3.5 “地区管理”基本信息管理

在“基本信息管理”界面下，单击“地区管理”按钮，显示页面如图 26 所示，可查询地区及邮箱，点击“下一页”可翻页浏览。

<sup>91</sup> 部分二级学科名称由于太长，在系统里采用简写，但是学科代码与国务院学位委员会和国家教育委员会《授予博士、硕士学位和培养研究生的学科、专业目录》（1997 年颁布，人文社会科学部分）完全一致。



图 26 基本信息管理 - 地区管理界面

#### 4.4.3.6 “专题信息” 基本信息管理

在“基本信息管理”界面下，单击“专题信息”按钮，显示页面如图 27 所示，第一列为专题代码，第二列为专题名称，点击“下一页”可翻页浏览。

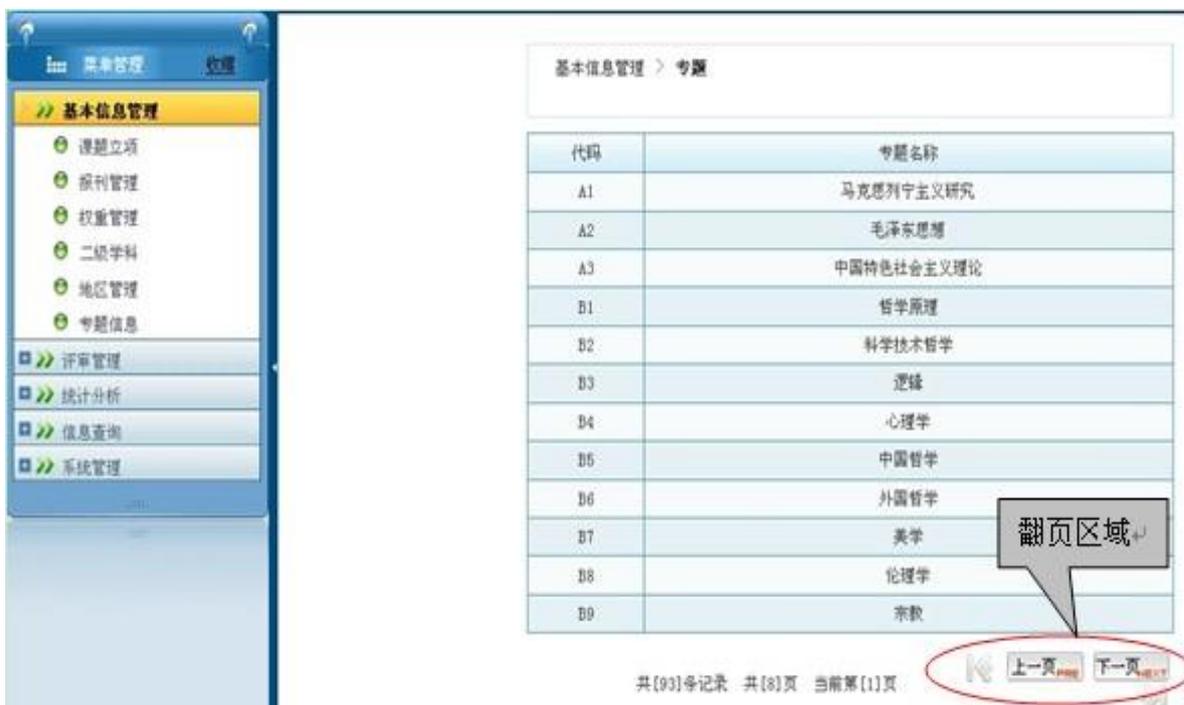


图 27 基本信息管理 - 专题信息管理界面

#### 4.4.4 评审管理

单击左侧导航栏里的“评审管理”按钮，进入评审管理界面，其功能包括：分类管理、评分管理、评分分布和评分明细。评审的基本流程为：先分类、后评分，评分分布和评分明细供评分时参考，评审流程图如图 28 所示。

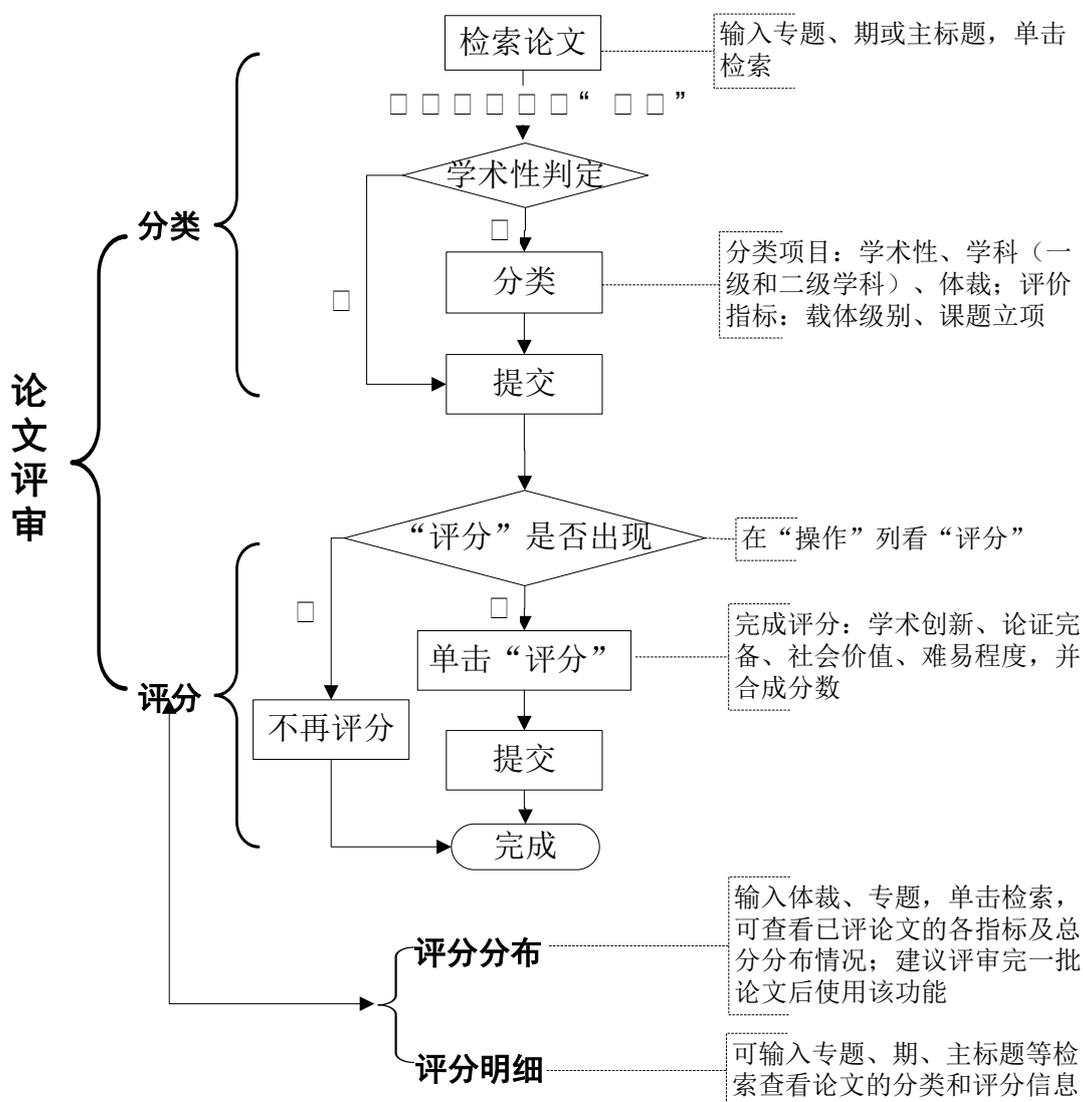


图 28 评审流程

#### 4.4.4.1 分类管理

##### (1) 功能简介

分类管理的功能是为论文的学术性、学科、体裁三项做分类标引，并判定课题立项、载体级别两个指标的级别。只有经过分类标引的论文才能进入评分环节。

##### (2) 使用操作

单击“评审管理”——“分类管理”，输入专题和期数，单击检索，进入分类管理界面，如图 29 所示，分类管理界面分为三部分：检索区域、显示区域（表格形式）、翻页区域（最下面一行）。检索区域的检索条件包括：专题（即专题代号）、期、作者、主标题、学术性（下拉菜单中行为全部、是、否）。“专题”字段支持精确匹配，须准确输入专题刊的代号（大小写均可）；“期”须准确输入 01、02、03、……、12，如输入 01 可检索出第 1 期的论文；“作者”和“主标题”字段支持模糊检索，只要被检索字段中包含检索词，就可将其全部检索到，而不用在意检索词在被检索字段中的位置。显示区域中可以浏览到论文题录信息和分类标引信息。

专题  期  作者  主标题  学术性

专题	期	主标题	副标题	作者	报刊名称	版别	一级学科	体裁	级别	课题立项	学术性	操作
A1	01	恩格斯晚年的观察与思考	文献档案的历史考证	徐宽哉	当代世界与		哲学	研究论文	核心	国家级	是	分类 评分
A1	01	关于“马克思主义理论”学科建设的几个问题	兼论新中国60年马克思主义理论研究中的经验	韦定广	南京政治学		哲学	研究论文	核心	省部级	是	分类 评分
A1	01	开启真理宝库的七把金钥匙	研读《共产党宣言》感悟七篇序言	高旻	中国浦东干		理论经济学	研究论文	核心	其他立项	是	分类 评分
A1	01	列宁关于小农改造的基本思想及其当代启示		彭海虹	社会主义研							分类
A1	01	列宁以两种思维方式阐发物质观的缘由之反转		贺祥林	江汉论坛							分类
A1	01	马克思《博士论文》与恩格斯《谢林和启示录》		鲁克俭	北京行政学							分类
A1	01	马克思公正思想维度下的全球公正问题解析		汪盛玉	社会主义研							分类
A1	01	马克思历史理论中的市民社会概念		阎孟伟	天津社会科							分类
A1	01	马克思主义理论学科建设支撑思想政治理论课		陈锡喜	思想理论教							分类
A1	01	马克思主义学习型政党建设的历史定位和理论	信仰、知识、方法的有机统一	程美东	马克思主义							分类
A1	01	西方马克思主义的理论性质与中国意义		刘同勋	中国社会科							分类
A1	02	“马克思主义”词义及其演变考察		林坚	马克思主义							分类
A1	02	从整体性角度把握马克思主义		张雷声	甘肃社会科							分类

共[13760]条记录 共[1059]页 当前第[1]页

图 29 分类管理

单击分类管理界面“操作”列中的“分类”，进入论文分类窗口。如图 30 所示。



图 30 分类管理——论文分类

论文分类窗口中，第一行为论文的题录信息：主标题和作者，题录信息下显示了需要做分类标引的项目（分类标引完成的项目前会显示红色的“ok”字符）。

➤ 学术性选择“是”表示该论文是学术文章，选择“否”表示该论文是非学术性文章。学术性选择“否”的文章，直接单击“提交”即可，不必再对文章的其他信息标引；学术性选择“是”的文章，需要继续对每个标引项目进行标引。

➤ 学科栏是指对论文的一级学科和二级学科信息作标引，单击论文所在的一级学科名称（只能单选），页面右侧立即弹出该一级学科下所含的二级学科信息，选择该论文归属的二级学科（可多选）。所有的一级学科和二级学科分类表，单击系统首页的“帮助”按钮查看。

➤ 体裁栏是判断论文是“综述文章”还是“研究论文”（单选）。

➤ 评价指标栏中的两项标引信息是评价论文的辅助指标，也是评价论文的客观指标，包括：载体级别和课题立项。其中，载体级别由系统判断，所以选项显示灰色且不可修改，无需评委选择。课题立项指标从四个相应的级别中选择一个（单选）。

➤ 所有标引项目完成后，单击“提交”。

论文分类后，学术性选择“是”的文章就会自动进入“评分”环节，选“否”的文章不进入“评分”环节，即无须评分。评分功能详见 4.4.4.2 部分的介绍。

### (3) 注意事项

(1) 在分类管理的界面, 操作列中的按钮若显示成黑色, 表示该操作未完成, 红色表示该操作已完成; 当分类标引完成后, “分类” 按钮变红; 若该论文是学术性文章, 则“分类” 按钮旁出现黑色的“评分” 按钮, 若该论文是非学术性文章, 则“分类” 按钮旁不会出现“评分” 按钮。

(2) 转载上的论文绝大部分都是学术性文章, 只有个别文章如卷首语是非学术性文章; 非学术性文章不进入后续的评分环节中, 因此在“评审管理-评分管理” 中检索不到非学术性文章。

(3) 在检索论文时, 专题号不区分大小写; 期号必须输入两位, 比如“01”; 若只填写专题号, 不填写期数, 则默认检索出该专题下所有期数的论文。本系统中其它界面的检索规则同此, 不再赘述。

(4) 显示项中内容显示不全的字段, 鼠标放在该单元格上, 可以显示该单元格的全部内容。本系统中其它界面的显示项显示不全的单元格也具有该功能, 不再赘述。

(5) 分类管理独立界面, 是为了支持今后的多评委对同一篇论文评价。当多评委进行评分时, 只需进行一次分类即可, 其他评委只对其他四个指标进行评分。

(6) 载体级别指标是由系统自动判定的。点击系统首页右上角的“帮助” 按钮, 可查看“载体级别” 文件。若评委发现某论文的原发报刊级别与该文件不符, 请与评价中心联系, 以免影响论文得分。

(7) 如图 30 所示, 论文分类界面上的“○” 表示单选, “□” 表示多选, 因此除了二级学科项目可多选外, 其余项目都只能单选。

#### 4.4.4.2 评分管理

##### (1) 功能简介

评分管理的功能是为论文的学术创新指标、论证完备指标、社会价值指标、难易程度指标评分, 并合成计算包括载体级别、课题立项在内的六个指标评分的

总分。评分管理是整个评估系统的核心功能。

## (2) 使用操作

单击“评审管理”——“评分管理”，输入专题和期数，单击检索，进入评分管理界面，与分类管理界面类似，此处略去。

单击操作列中的“评分”按钮，进入论文评分界面，如图 7 所示，第一行方框中显示论文的题录信息：主标题和作者，题录信息下显示四个主要评价指标：学术创新、论证完备、社会价值、难易程度，鼠标放在指标名称上还可看到该指标的内容解释。四个指标的评分尺度采用 21 分量表打分法：

- 每一个“○”代表 1-21 间的一个值，将鼠标移动到“○”上时就可看到该位置“○”代表的分数，从左到右依次减小，最大值为 21 分，最小值为 1 分；
- 评委根据自己对论文在某方面表现的印象，单击量表上相应位置的“○”即可，右侧“信息显示”栏会显示当前指标的得分，同时指标前出现红色的“ok”字符；
- 评分尺上标注了极好（21 分）、较好（16 分）、一般（11 分）、较差（6 分）、极差（1 分），当某篇论文在某个指标上表现极好，则在量尺上“极好”的位置（21 分）单击“○”；当某篇论文在某个指标上处于“极好”和“较好”之间，则评委在量尺上“极好”和“较好”的位置之间根据印象再选择合适的评分。

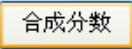
所有指标评分完成后，先单击“”按钮，在页面右侧的“信息显示”栏会显示总分；如果评分过程中有指标未评分就单击“合成分数”，则页面会提示“XXXX 没有选择”，表示合成分数失败，评委需要将遗漏指标的评分补充完整，重新单击“合成分数”。分数合成后，若“信息显示”——“最后得分：”符合预想总分，再单击“”按钮，则对该篇论文的评审结束；若不符合预想总分，可重新评审并再次合成分数直至符合预想的总分，并单击“”，完成评分。

图 31 评分管理——论文评分

### (3) 注意事项

(1) 页面上有两个进入“论文评分”的入口：一是单击分类管理界面的“操作”列中出现的“评分”按钮，弹出“论文评分”界面；二是在导航栏中“评审管理——评分管理——显示项中的操作列——单击评分”进入评分界面。建议评委采用第一种进入“评分”入口的方式，分类后直接进入论文评分界面。

(2) 评审过程中遵循“先分类，后评分”的程序，只有完成分类的学术性文章才可评分。

(3) 修改“论文评分”界面的评分时，评委只需修改其中个别评分并合成分数——提交，而不需要对全部指标重新评分。

(4) 若论文已经完成了评分，但又要对分类信息进行修改时，系统会自动根据最新的分类信息，重新合成计算该论文评分的总分，因此不需要重新评分。

(5) 总分是学术创新、论证完备、社会价值、难易程度、载体级别、课题立项六项指标得分按学科和体裁所设置的权重求和的结果。指标权重分配方案详见本系统页面右上角的“帮助”-“权重分配方案”。论文总分计算公式如下，其中 $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6$ 是权重：

总分=0.9×(λ<sub>1</sub>×学术创新程度指标评分+λ<sub>2</sub>×论证完备程度指标评分+λ<sub>3</sub>×社会价值指标评分+λ<sub>4</sub>×难易程度指标评分)+0.1×(λ<sub>5</sub>×载体级别指标评分+λ<sub>6</sub>×课题立项指标评分)

注：辅助指标的权重由去年的 0.15 调整为 0.1。

(6) 评分管理模块支持多评委对同一篇论文的评分。但 2011 年度转载论文只由一名评委评分，因此不要评别人负责的专题。

### 4.4.4.3 评分分布

#### (1) 功能简介

评分分布的功能是帮助评委在评审过程中随时掌握已评论文的评分分布情况，矫正评委评分尺度过松或过严的极端情况，避免主观性造成的评分误差。该功能是通过统计某专题下、某评委已评论文的各指标评分出现频次和比例、平均值、分数覆盖数量、总分分布等指标，并将评分分布不合理的标红显示。

#### (2) 使用操作

单击“评审管理——评分分布”，系统弹出“评分分布查询”界面，第一行的检索框有三项信息：评分人（默认登录系统的当前评委）、体裁（默认研究论文）、专题（默认空），评委在“专题”检索框中输入专题代号，单击“检索”按钮，即可检索出该评委在该专题下的研究论文的评分情况，如图 32 所示。



图 32 评分分布查询

评分分布查询界面中用两个表格来展示评分统计情况：各指标评分分布表和

总分分布表。

➤ 图 32 的第一个表中，首列“分数”列是评分值，由于采用了 21 分量表，因此该表中最多可显示 21 个分值。

➤ “分数”列右侧的“学术创新指标”列统计了该指标在每一评分分值上出现的频次和比例，“学术创新指标”列与“平均得分”行交叉的“17.0”是该专题下，该评委在已评研究论文中的学术性创新指标评分的平均值，即已评的研究论文的学术创新指标评分的平均值为 17.0。

➤ “分数覆盖数量”是指覆盖了多少个不同评分分值，如“分数覆盖数量”行与“学术创新指标”列交叉的“2”表示学术创新指标评分覆盖了 2 个不同评分，即 9 分（1 个）、21 分（2 个）。

➤ “论文数量及比例”一行统计了该专题下当前已评研究论文的数量占该专题已评论文数量的百分比，以图 32 中的“3/4(75.0%)”为例，“3”表示该专题下当前已评研究论文数量，“4”表示该专题下已评论文数量，3 除以 4 表示已评研究论文占已评所有论文的比例，即 75.0%。

➤ 表下的文字“单项指标得分分布参考值……”是评分分布的理想状况，如果评分分布结果不符合此参考值，将在单元格中以“红色”显示。

➤ 在第二个表中，第一行是总分的分数区间段，共有 5 个；频次是指每个区间段中总分出现的频次，如“[21,18)”与“频次”交叉处的“2”表示在已评研究论文中，有 2 篇论文的总分处于[21,18)；比例是指研究论文中，各区间段中评分出现的频次占已评研究论文数量的比例，如“[21,18)”与“比例 (%)”交叉处的“66.67”表示总分处于[21,18)的文章数量占已评研究论文数量的比例。

### (3) 注意事项

(1) 系统中进入“评分分布”有两个入口：一是从菜单导航栏“评审管理——评分分布”进入，二是从分类管理界面或评分管理界面右上角的“评分分布查询”进入。此外，点击“评分分布查询”按钮后，需要再次输入专题等信息，并单击“检索”才能出现评分分布结果。

(2) 评分分布可动态统计, 在使用时最好是评完 5 篇以上的论文再使用该功能, 因为已评论文数量过少统计评分分布无意义。

(3) 检索框中“评分人”默认是登录系统的评委姓名, 若在“评分人”检索框中输入其它评委姓名, 也可查看其它评委评分分布情况。

(4) 系统将单项指标的评分分布不符合“参考值”的地方标红显示, 仅供评委参考, 以便保证对论文的评分具有较好的区分度: 若评分分布表中出现标红的信息, 评委应根据此信息到“评分明细查询”界面, 按各指标和总分的区间段查询相应评分段的论文, 并修改该评分; 若评委若认为所评论文的质量确实如此, 也可忽略该标红信息, 不再修改评分。

#### 4.4.4.4 评分明细

##### (1) 功能简介

该页面功能用来检索、显示和导出单篇论文的原始分类信息和评分情况, 包括: 一级学科、二级学科、体裁、载体级别、课题立项、学术创新指标、论证完备指标、社会价值指标、难易程度指标、总分、评委和评分时间。

##### (2) 使用操作

单击“评审管理”——“评分明细”, 输入专题和期数, 单击检索, 进入明细查询界面, 如图 33 所示。界面分为检索区域、显示区域和翻页区域, 检索区域中的检索条件有: 专题、期、主标题、评委、指标项、分数区间, 支持多个检索条件联合检索。其中, “专题”检索框的有效输入值是专题代号, 如 A1; “期”检索框的有效输入值是 01、02、……、12; 检索主标题检索框支持模糊检索; 指标项为下拉菜单, 包含: 学术创新、论证完备、社会价值、难易程度、总分; 分数区间“从 到 ”中可输入 1-21 分之间的任何整数。该页面右侧的“导出”按钮还可以 EXCEL 形式导出论文评分明细。

专题	期	主标题	副标题	报刊名称	版别	一级学科	二级学科	体裁	载体	课题	创新	论证	价值	难易	总分	评委	评分时间
A1	01	恩格斯晚年的观察与思考		文献档案	当代世界	哲学	[伦理学]	[研究论	21	21	20	20	20	20	20.1000	系统管理	2012-02-01
A1	01	恩格斯晚年的观察与思考		文献档案	当代世界	哲学	[伦理学]	[研究论	21	21	20	20	20	20	10.9390	刘峰	2012-02-02
G9	01	《公共图书馆法》立法进展		图书馆建设		公共管理	[教育经济研究论		21	1	15	16	16	15	15.1910	杨红艳	2012-02-01
G9	01	ClimateQUAL图书馆服务质量评价新体系		大学图书馆		图书馆	[图书馆学研究论		21	8	11	10	10	11	10.9390	杨红艳	2012-02-01
G9	01	NSTL联机联合编目工作的发展与实践		数字图书馆		图书馆	[图书馆学研究论		21	8	14	17	17	14	15.4390	杨红艳	2012-02-01

共[5]条记录 共[1]页 当前第[1]页 << 上一页 PRE 下一页 NEXT >>

图 33 明细查询

### (3) 注意事项

1) 多评委的评分结果分开展示，如相同专题下同一篇文章由两位以上（含两位）评委评审，则系统将分别展示每位评委评分结果，如图 33 中前两行为 A1 转载的同一篇文章，分别由两位不同评委评审，则系统分别展示两位评委的评分结果。

2) 由于学术性为“否”的文章不进入评分环节，因此明细查询中查不到该类文章。

3) 由于网页空间有限，表中部分标题采用简写形式，载体（即载体级别），课题（即课题立项），创新（即学术创新），论证（即论证完备），价值（即社会价值），难易（即难易程度）。

4) 由于被转载论文水平较高，理论上其总分不应低于 11 分；而实际评分结果出现低于 11 分的总分时，以红色显示，提示评委该论文评分过低，建议修正。

## 4.4.5 统计分析

单击左侧导航栏里的“统计分析”按钮，进入统计分析页面，统计分析有归一计算、统计报表两个功能模块，本指南仅介绍统计报表功能模块。

### 4.4.5.1 统计报表

“统计报表”功能模块提供期刊排名、作者单位排名的报表查询与导出功能。

#### (1) 查询功能

进入“统计报表”页面，导航栏左侧以树状结构显示各期刊排名情况，当前显示“+”，表示该类下还可点击子类，当前显示“-”，表示该类下的子类已经全

部展开。

如图 34 所示，可按照报刊名称、报刊内容、出版地三个条件查找论文。报刊内容下拉菜单可选择“全部”、“综合性”、“非综合性”等限定条件，出版地下拉菜单可选择相应地区。单击“查询”，即可检索到相关报刊的基本信息和得分情况。

报刊名称	版别	刊期	页码	地区	报刊内容	综合得分	学术创新	论证完备	社会价值
《北京大学学报》	哲学社...	双月	160	北京	综合性	30.872933	34.269046	34.345240	33.161
《北京师范大学学报》	社会科...	双月	144	北京	综合性	20.994175	22.692858	23.392860	22.841
《财经科学》		月刊	124	四川	非综合性	24.125205	26.140480	26.078065	26.881
《安徽大学学报》	哲学社...	双月	156	安徽	综合性	14.316345	15.714287	16.071429	15.221
《北京交通大学学报》	社会科...	季刊	136	北京	综合性	7.029550	7.325313	7.752507	8.005
《北京航空航天大学学报》	社会科...	双月	120	北京	综合性	7.169904	7.895239	7.597619	7.876
《东北师大学报》	哲学社...	双月	236	吉林	综合性	8.233890	8.652380	8.990476	9.092
《北京体育大学学报》		月刊	144	北京	非综合性	8.809096	9.809525	9.761906	9.285
《东南大学学报》	哲学社...	双月	128	江苏	综合性	5.566125	5.933334	6.166666	6.113
《东北大学学报》	社会科...	双月	94	辽宁	综合性	4.771119	5.326191	5.094737	5.130
《大连理工大学学报》	社会科...	季刊	128	辽宁	综合性	3.255099	3.373684	3.550003	3.509
《电子科技大学学报》	社科版	双月	112	四川	综合性	3.401548	3.664285	3.711905	3.664
《北京工业大学学报》	社会科...	双月	80	北京	综合性	3.391260	3.709525	3.759524	3.709
《北京理工大学学报》	社会科...	双月	156	北京	综合性	2.452410	2.598997	2.602381	2.697
《北京邮电大学学报》	社会科...	双月	116	北京	综合性	2.939713	3.056639	3.361907	3.171

图 34 统计分析 - 统计报表查询

## (2) 报表导出功能

查询完成后，单击页面右上角的“导出”按钮，即可导出 Excel 格式的报表。

### 4.4.6 信息查询

提供对所有已评论文的查询情况。评委可看到本人及他人负责专题的评分进度。包括对各评委&专题论文数查询、评委论文数查询、专题论文数查询、每日评论文数的查询。如图 35 所示：

评委&专题论文数查询		
评委	专题号	已评论文数
刘峰	A1	2
系统管...	A1	3
杨红艳	A1	3
杨红艳	C1	4
范晓莉	G9	1
杨红艳	G9	3
合计		16

评委论文数查询表	
评委	已评论文数
范晓莉	1
刘峰	2
系统管理员	3
杨红艳	10
合计	16

专题论文数查询		
专题号	已评论文数	待评论文数
A1	5	133
A2	0	59
A3	0	191
B1	0	226
B2	0	133
B3	0	45

每日评论文数查询	
日期	已评论文数
2012-02-01	4
2012-02-02	10
2012-02-03	2
合计	16

图 35 信息查询

- “评委&专题论文数查询”统计了某评委在某专题里评审的论文数。
- “评委论文数查询”统计了评委跨专题评审的论文数。
- “专题论文数查询”统计了各专题里的已评论文数和待评论文数。多评委同评一个专题时，已评论文数不进行重复统计。比如 3 名评委对同一篇论文进行了评分，已评论文数只计为 1 篇。
- “每日评论文数查询”统计了所有专题的每日评审论文数，不是针对某一专题的统计。

#### 4.4.7 帮助功能

提供评委在系统使用过程中，可能需要的各类信息。单击页面右上角的“帮助”按钮，评委可查看与评审相关的文件，包括：

- 指标含义及内容
- 发表载体级别

- 指标权重分配
- 论文学科分类表
- 评审操作指南