

学位与研究生教育学科评估研究^{【1】}

北京航空航天大学 魏晖、郑晓齐、韩慧瑜

一、前言

从2002年开始,教育部学位与研究生教育发展中心开展了一级学科整体水平评估试点,这是一件非常有意义的工作,也是我国研究生教育评估中出现的新事物,备受各界的关注。有关一级学科整体水平评估的理论、指导思想、原则和具体方法等有很多问题值得研究和探讨。本研究通过查阅资料,对英国、美国、中国有关学科评估情况进行分析,选取中国的一级学科整体水平评估、英国的科研评估和美国1976-1982年研究型博士学位点评估进行比较分析,重点对组织者、评估单元、评估指标体系、评估结果、评估周期、评估主要方式等进行比较,在此基础上提出一级学科整体水平评估的原则和有关建议。

二、英国开展研究生教育学科评估情况简介^{【2】【3】【4】}

英国高等教育评估中影响最大的是科研评估和教育质量评估。前者由英格兰、苏格兰、威尔士3个高教拨款委员会和北爱尔兰教育部联合组织,已分别于1986,1989,1992和1996年进行了4次全英范围内的科研评估,主要评估博士生和研究型硕士生教育。后者由英格兰和苏格兰高教委员会分别组织,不存在全英范围内统一的教育质量评估,主要评估本科和课程型硕士生教育。分别开展科研和教育质量评估是为了突出高等学校人才培养的重要性。他们认为重科研轻教学,甚至以牺牲教学特别是本科教学来保全科研是全世界高等学校的通病,如果将科研和教育质量放在一起评估,会使公众对某些科研水平高而教育质量一般的大学产生误解。科研和教育质量评估的具体情况见表1。

【1】谢桂华,教育部学位与研究生教育发展中心“十五”课题研究成果汇编——学位与研究生教育研究新进展,高等教育出版社,380-391,2006.6

【2】中国高等学校学科评估代表团.对英国高教评估与拨款制度的考察.学位与研究生教育,1998(3)

【3】陈钟欣.英国高校的教育质量评估.中国高等教育评估,1999(1)

【4】杨于文.改革学科评估,促进学科发展.中国高教研究,1999(3)

表 1 英国高等教育科研和教育质量评估活动简表

评估类型	科研评估 (RAE)	教育质量评估 (QA 或 AQE)	
组织者	英格兰, 苏格兰, 威尔士三个高教拨款委员会 (HEFCE, SHEFC, HEFCW) 和北爱尔兰教育部 (DENI) 联合组织进行	英格兰高教拨款委员会 (HEFCE)	苏格兰高教拨款委员会 (SHEFC)
被评对象	全英国所有主动申请参加评估的高校各学科	英格兰高校等	苏格兰高校
评估周期	3 或 4 年一次 (已经进行 4 次: 1986, 1989, 1992 和 1996)	各评估单元轮流进行, 8 年为一周期	各评估单元轮流进行, 6 年为一周期
评估单元	1996 按学科分为 69 个单元 (UOA)	按学科分为 61 个单元 (UOA)	按学科划分评估单元 (cognate areas), 可结合学校情况较灵活地确定
主要方式	同行评议为主, 如 1996 成立了 60 个学科组 (panel) 评议所有 69 个单元	在自评基础上同行专家评估, 一定要到校实地考察	在自评基础上同行专家评估, 一定要到校实地考察
指标体系	各学科组 (panel) 自定并提前公布本组用的评估指标内容, 一般重点考察: 论著及其它成果*; 外来科研收入; 从事研究的学生数及获研究奖学金人数; 科研规划与发展前景等。只有少数专家组为各一级指标定出权重系数范围。	包括 6 大方面: 1. 课程计划; 2. 教学与评估; 3. 学生进展与成绩; 4. 对学生的支持与指导; 5. 学习资源; 6. 质量保证和提高	包括 12 项指标 (略)
评分等级	1989 和 1992 年按 1-5 分共 5 级赋分值; 1996 年改为按 1, 2, 3b, 3a, 4, 5, 5* 共 7 级赋分, 各学科点得分与被评估科研人员数 (比例) 同时公布	按 1, 2, 3, 4 分 4 级赋值。1 分为不合格	按优秀、高度满意, 满意和不满意 4 级评定
评估的其他特色	鼓励竞争, 评估结果与高教拨款委员会下拨科研经费挂钩, 既影响声誉, 又影响科研经费	重在保证教学质量, 评估结果尚未与拨款挂钩	重在保证教学质量, 评估结果尚未与拨款挂钩

*1992 年前考察论著总数, 导致片面追求发表论文篇数而不利于提高科研水平和效益。从而, 1996 年改为考察每位科研人员 (research active staff) 提供的 4 项最代表本人水平的成果。

(一) 评估单元

英国科研评估一律以学科单元为基础, 不对学校整体科研工作进行评估。学科单元是在分析了英国各高校现有的研究方向, 并广泛征询专家意见的基础上确定的, 它和我国的一级学科、二级学科有所不同, 每一学科单元包括的内容一般比我国的一级学科、二级学科要广泛。划分的 69 个学科单元分别为:

1. 临床实验室科学; 2. 公共临床医学; 3. 医院临床学; 4. 牙医学; 5.

诊断前研究(预诊研究); 6. 解剖学; 7. 生理学; 8. 药理学; 9. 药学(制药); 10. 护理学; 11. 其他有关药物类研究; 12. 生物化学; 13. 心理学; 14. 生理科学; 15. 农业; 16. 食物科学与技术; 17. 兽医学; 18. 化学; 19. 物理; 20. 地球科学; 21. 环境科学; 22. 纯数学; 23. 应用数学; 24. 统计与运筹学研究; 25. 计算机科学; 26. 通用工程; 27. 化学工程; 28. 土木工程; 29. 电气与电子工程; 30. 机械、航空与制造工程; 31. 矿冶工程; 32. 金属和材料; 33. 环境结构; 34. 城市规划; 35. 地理学; 36. 法律; 37. 人类学; 38. 经济学; 39. 国际政治学研究; 40. 社会政策和管理; 41. 社会工作; 42. 社会学; 43. 商业和管理研究; 44. 会计学; 45. 美洲研究; 46. 中东和非洲研究; 47. 亚洲研究; 48. 欧洲研究; 49. 凯尔特语研究; 50. 英语和文学; 51. 法语; 52. 德语、荷兰语和斯堪的拉维亚语言; 53. 意大利语; 54. 俄语、斯拉夫语和东欧语言; 55. 爱伯瑞和拉丁美洲语言; 56. 语言学; 57. 经典、古代历史、拜占庭和现代希腊语研究; 58. 考古学; 59. 历史学; 60. 艺术、建筑、设计史; 61. 图书和信息管理; 62. 哲学; 63. 神学、神性和宗教研究; 64. 设计艺术; 65. 交流、文化、媒介研究; 66. 戏剧、舞蹈、表演艺术; 67. 音乐; 68. 教育学; 69. 体育运动相关研究。

(二) 评分标准

评分标准分为 5*、5、4、3a、3b、2、1 等 7 级，每一级代表的水平分别为：

5*——大多数学科子领域为国际先进水平，其它都达到国内先进水平；

5——某些学科子领域达到国际先进水平，其它均达到国内先进水平；

4——全部学科子领域为国内先进水平，可能某些达到国际先进水平；或者某些达到国际先进水平，大多数为国内先进水平；

3a——相当大多数学科子领域达到国内先进水平，或达国际先进水平和国内先进水平二者之和构成了大多数学科子领域；

3b——大多数学科子领域达国内先进水平；

2——半数学学科子领域达国内先进水平；

1——各学科子领域均未达国内先进水平。

(三) 评估方法

1996 年完成的科研评估方法主要包括：

建立 60 个学科专家组以负责 69 个评估单元中的所有评估工作。

各学校按评估单元准备和提交书面报告，其形式对于所有学科相同，内容包括：

①1996 年 3 月 31 日仍在工作的待评估科研人员名单(各评估单位自行确定所报人员名单，不一定包括拥有的全部人员)；

②在给定期间（某些文科为6年，其他学科均为4年）每一位科研人员的科研成果（4篇论著或其它研究成果）；

③在所评估期间从事研究的学生及享受研究奖学金的人数、科研收入；

④关于本系的研究成就、支持与发展科研的安排、外界认可的标志等的说明，以及学校希望能引起学科专家组注意的其它信息。

各校自行决定各系及有关研究人员向哪一个评估学科组申报，并允许两个学校的学科联合申报。

在各校提交书面报告截止期的6个月前，各学科专家组已经开会确定了各自的评估指标体系及工作方法，并与提交书面报告的详细指导书一并公布。

1996年4月30日提交书面报告截止后，学科专家组在5-11月间根据原已公布的评估体系进行评估。必要时，征求其它学科专家组或外界专家的意见。

各校提交的书面报告均经过多种方式的核查，并最后按7级分别打分。公布时除发布评价分值外，还要公布参加评估的教师比率。只有在得到5*的同时，参加人员比率最高，才算是最好的学科。牛津、剑桥除5*学科较多外，参加评估的教师接近教师总数的100%，这也是其声誉和水平较高的标志之一。

三、美国开展研究生教育学科评估情况简介【5】【6】【7】【8】【9】

美国对研究生教育评估始于1924年，1934年和1957年又进行过两次。这三次评估都是由一所大学组织的，规模较小，其基本目的只是弄清该校在水平较高的大学中的地位，评估结果由美国教育理事会出版。1964年由美国教育理事会出面组织，对106所大学中的1633个授予博士学位的学科点进行了评估，主持者仔细研究了公众对以往几次研究生教育质量评估的反映，在作法上做了改进，又以几个学科为典型对评估结果的可靠性进行了检验。这次评估可以说是美国对研究生教育正规评估的开始。1966年Allan M. Cartter发表了《关于研究生教育质量的评估》的研究报告。1969年，美国教育理事会又对130所大学中的2626个授予博士学位的学科点进行了评估。1970年D. Roose及C. J. Andersen发表了《研究生学位点评估》的研究报告。1976年-1982年美国教育理事会等4家全国性的学术、教育组织组成联合研究委员会，领导了美国研究型博士学位学科点的评估。这次评估的经费由梅隆基金、福特基金、斯隆基金、美

【5】刘盛钢. 研究生学科的评估. 美国加拿大高等教育评估第四分册. 成都: 成都电讯工程学院出版社, 1987

【6】陈钟欣. 中美“研究生院评估”比较. 学位与研究生教育, 1996(4)

【7】沈红. 美国研究生教育评估的第一层: 分学科门类的评估. 学位与研究生教育, 1996(4)

【8】沈红. 美国研究生教育评估的第二层: 分学科级别的评估. 学位与研究生教育, 1996(5)

【9】曲铭峰. 美国最好的研究生院---《美国新闻和世界报道》对美国各大学研究生教育的评估. 学位与研究生教育, 2001(2-3)

国卫生组织（NIH）、国家科学基金会（NSF）和美国科学院（NAS）提供。美国地理学会、美国心理学会和美国心理学基金会也为该研究提供了有效的财政资助。1976年9月开始制定评估方案，1978年9月，计划委员会提供了一个可以对质量评估和研究型博士学位点有效性描述的策略性的综合报告。这份报告构成了从事科学、工程和人文科学方面博士学位点国家级评估的会议委员会决议的基本内容。1980年6月，任命了一个18人委员会，具体主持这项调查研究工作。该委员会由自然科学、工程学及人文科学范围内完全不同类别的学科的人员组成。其中的7名委员，曾参加制定计划，另外几个成员在大学担任或曾经担任过研究生院院长。

这次参与评估的有228所大学，分成31类学科，包括2699个研究生学位点。这次评估的最大特点在于建立了比较完善的评估指标体系，改变了过去单纯采用“主观评估”（也称“声誉评估”、“同行评估”或“社会评估”）的办法。

这次评估共分为31类学科，分别为：

艺术和人文科学：艺术史，古典文学，英国语言文学，法兰西语言文学，德国语言文学，语言学，音乐，哲学，西班牙及葡萄牙语言文学；

生物科学：生物化学，植物学（含植物生理学、病理学、真菌学），细胞生物学/分子生物学，微生物学（含免疫学、细菌学、寄生虫学、病毒学），生理学（动物、人类），动物学；

工程学：化学工程，土木工程，电气工程，机械工程；

自然科学 化学，计算机科学，地球科学（含地质学、地球化学、地球物理学、普通地球科学），数学，物理学（不含天文学、天体物理学），统计学；

社会科学和行为科学：人类学，经济学，历史学，政治科学，心理学，社会学。

计划委员会和评估委员会都仔细研究了用于这次研究型博士学位点评估的指标体系。两个委员会一致认为，任何可供使用的指标都有不足之处，而且不可能期望能使用任何一个指标来给出有关研究生教育质量之完全满意的评价。在选择使用哪些指标时，主要考虑以下两个因素：①该指标与博士学位点质量相关联的程度；②能够进行各类专业学位点的全国性比较，编制可靠数据的可行性。

在工程学科博士学位点评估中选取了6个一级指标和16个二级指标：

（1）学位点规模

01 学位点教师数目（1980年12月为准）；

02 最近5年中学位点授予的博士学位数目（1975年7月-1980年6月）；

03 全部时间攻读和部分时间攻读的在校博士研究生总数（以1980年12月的注册数为准）。

(2) 毕业生的情况

04 研究生期间, 获得某种国家奖学金或教育基金资助的研究生比例(1975年-1979年财政年度);

05 1975-1979 财政年度期间, 研究生从其首次入学到获得博士学位所花的平均年数;

06 1975-1979 财政年度期间, 学位点研究生完成博士学位要求时, 已明确承诺了毕业后工作的研究生比例;

07 1975-1979 财政年度期间, 学位点研究生完成博士学位要求时, 已明确承诺了在可授予博士学位的大学工作的比例。

(3) 声誉调查结果

08 学位点教师学术水平的评估均值;

09 在培养研究学者/科学家方面, 学位点有效性的评估均值;

10 在最近 5 年中, 关于学位点质量改善的评估均值;

11 评估人对被评学位点教师工作的熟悉程度的评估均值。

(4) 大学图书馆规模

12 描述学位点所属大学图书馆规模的综合指数(以 1979-1980 年间数据为准)。

(5) 研究资助

13 在 1978-1980 财政年度期间, 从国家科学基金会, 国家健康研究院或酒精、药物滥用、精神病管理局取得研究资助的学位点教师成员的比例;

14 在 1979 年用于某一特定领域的研究与发展的总经费(该数目由学位点所在大学呈报并以千美元计)。

(6) 论文、著作统计

15 1978-1979 年度间发表的与学位计划有关的论文数目;

16 1978-1979 年度间发表的与学位计划有关的论文“总影响力”的估计。

自 20 世纪 80 年代后期起,《美国新闻与世界报道》对美国研究生教育每年进行一次全国性评估, 评选出美国最佳研究生院和专业学院(America's Best Graduate and Professional Schools) 和最佳专业学位点(Best Specialty Programs), 已成为影响力最大的评估。

评估的方法按学科不同而有所差异, 对授予 Ph.D 即哲学博士的学科, 主要采取同行评估的方式, 由特定学科领域中的研究生院院长们和教授们对各大学在该领域研究生教育的学术声誉评分, 最高分为 5 分, 然后根据各大学所获同行专家评分的平均值按大小排序。对授予具有明确职业定向的专业学位的学科, 则采取综合评估的方式,

除了同行评估外，还要进行社会评估，即对某一特定学科领域，都要征询学术圈外有关专家的意见，请他们对各大学在该学科领域研究生教育的学术声誉评分；同时还参照若干客观标准，包括研究经费、师资力量、学生就业情况、学生考试成绩和学生入学时对学校的选择顺序等，在综合这些指标的基础上，然后给出各大学在该学科领域研究生教育水平的总评分（其满分为 100 分），并按大小进行排序。

评选最佳研究生院和专业学院主要分法律、商业、工程、医学和教育 5 个学科门类，评估指标体系一般包括 9-10 项指标。如 1999 年“美国最佳工学院”指标为：学术声誉排名；用人单位给出的声誉排名；1998 入学考试平均成绩（定量、分析）；1998 录取率；博士生/教师；1998 教授中国家工程院院士之比；1998 工程研究开支；98 教授人均研究开支；1997-1998 博士授予数。

四、我国开展学位与研究生教育学科评估的情况

我国从 1985 年 2 月国务院学位委员会第六次会议决定逐步建立各级学位授予质量的检查和评估制度，18 年来，学位与研究生教育评估得到了迅速发展。国务院学位委员会办公室设立了“质量监督与信息工作处”，负责全国学位与研究生教育评估的组织与领导工作，并着手制定有关法规和政策；成立了“高等学校与科研院所学位与研究生教育评估所”，受国务院学位委员会和教育部的委托，承担有关学位与研究生教育的评估和咨询工作。已经开展的有关学位与研究生教育学科评估主要有增列博士、硕士学位授权点评估、一级学科选优评估、博士硕士学位授权点基本条件合格评估、高等学校重点学科评估和一级学科整体水平评估等。

增列博士、硕士学位授权点评估和博士硕士学位授权点基本条件合格评估是一种鉴定评估，也称为合格评估。一级学科选优评估、重点学科评估和一级学科整体水平评估等是一种水平评估，也称为选优评估。前面介绍的英国、美国学科评估都是水平评估。为了便于与英、美进行比较，下面重点介绍一级学科整体水平评估的情况。

2002 年 4 月开始的一级学科整体水平评估试点工作，选择了 12 个基础较好、学科授权点较多的一级学科，本次评估采取自愿参加的方式。开展一级学科整体水平评估的 12 个学科见表 2。

表 2 参加一级学科整体水平评估的学科范围

学科门类	学科名称
理科	数学
	物理
	化学
工科	机械工程
	电子科学与技术
	信息与通信工程
	控制科学与工程
	计算机科学与技术
	生物学工程
农科	作物学
	植物保护
管理学	管理科学与工程

考虑到学科之间的差异，评估指标体系及权重集分理工农类（见图 1）和管理类（见图 2）。

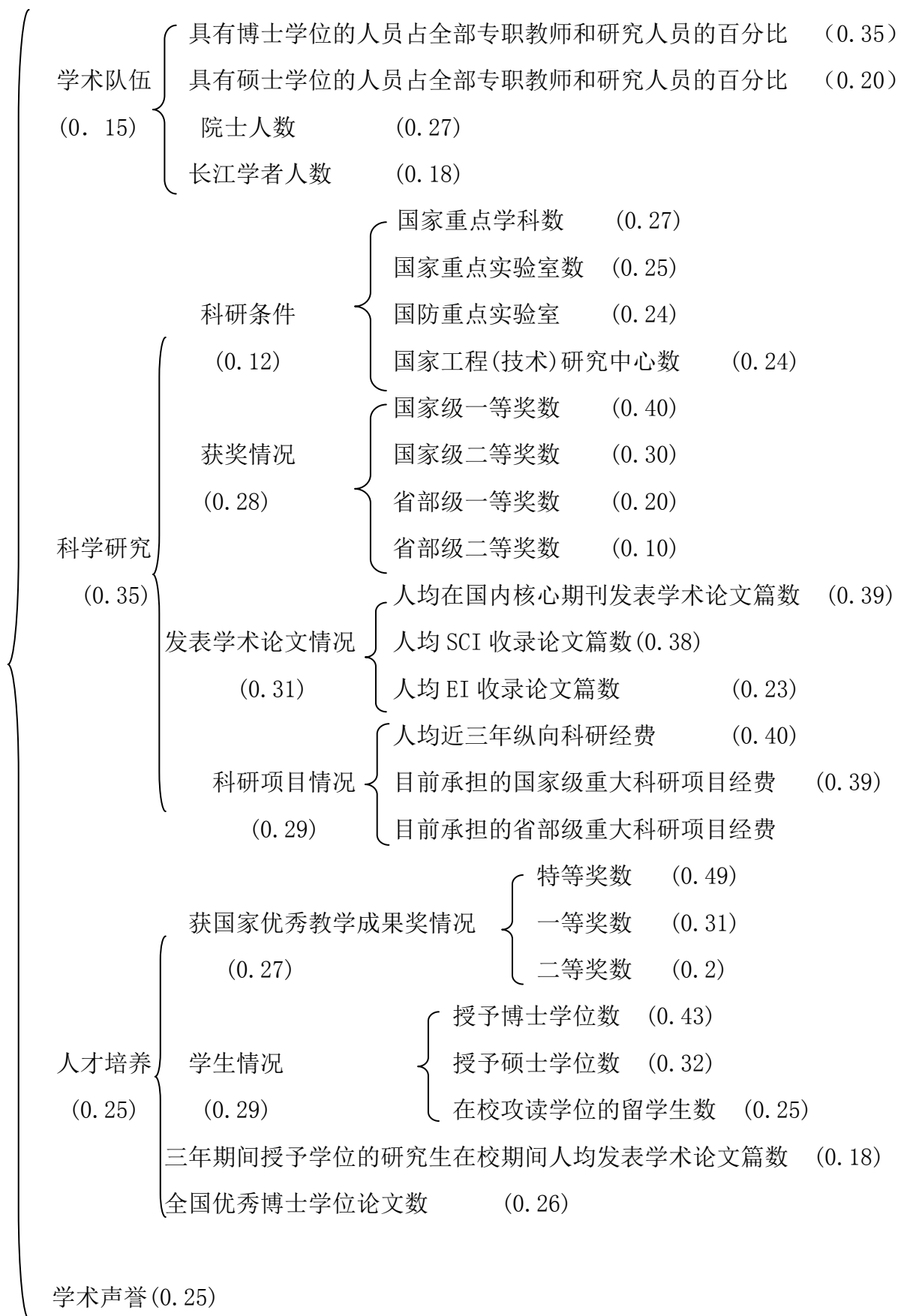


图 1. 一级学科整体水平评估指标体系及权重集(理工农类)

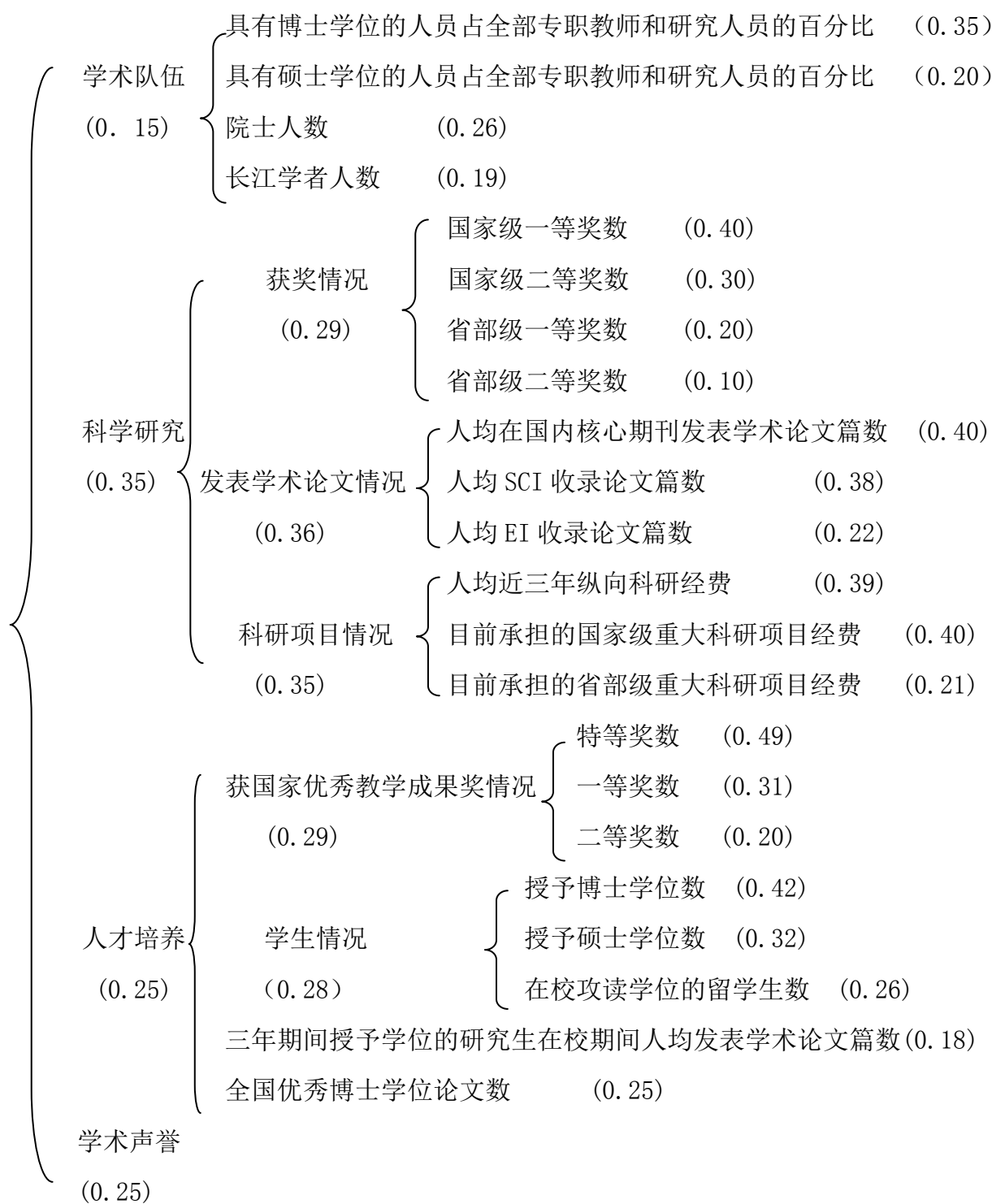


图2 一级学科整体水平评估指标体系及权重集(管理类)

五、中、英、美三国学科评估比较分析

为便于比较,选取中国的一级学科整体水平评估、英国的科研评估和美国1976-1982年研究型博士学位点评估进行比较分析。三国学科评估的比较见表3。

表 3 中、英、美三国学科评估比较分析

	中国	英国	美国
组织者	教育部学位与研究生教育发展中心	英格兰，苏格兰，威尔士三个高教拨款委员会 (HEFCE, SHEFC, HEFCW) 和北爱尔兰教育部 (DENI) 联合组织进行	美国教育理事会。
评估单元	一级学科	学科单元 (UOA), 共分 69 个单元	分成 31 类学科
指标体系	4 个一级指标: 学术队伍; 科学研究; 人才培养; 学术声誉。 每个一级指标都有相应的权重	各学科组 (panel) 自定评估指标内容, 一般重点考察: 论著及其它成果; 外来科研收入; 从事研究的学生数及获研究奖学金人数; 科研规划与发展前景等。只有少数专家组为各一级指标定出权重系数范围	6 个一级指标: 学位点规模; 毕业生情况; 声誉调查结果; 大学图书馆规模; 研究资助; 论文、著作统计。 16 个二级指标, 但没有权重
评估结果	分 4 个一级指标分别打分, 根据权重系数得出总分, 并根据总分和 4 个分项排序	分 7 级赋分, 各学科点得分与被评估科研人员数 (比例) 同时公布。评估结果与高教拨款委员会下拨科研经费挂钩, 既影响声誉, 又影响科研经费	16 个二级指标的评估均值和标准差
评估周期	原则上一年一次	3 或 4 年一次	不定期
主要方式	主观评估与客观评估相结合, 主观评估以学术声誉调查为主	以主观评估为主, 需要提供一些客观数据	主观评估与客观评估相结合, 主观评估以学术声誉调查为主

中国、英国、美国 3 国的评估各具特色, 中国的评估虽然历史较短, 但发展很快, 根据教育部颁发的授予硕士、博士学位和培养研究生的一级学科进行, 评估对象有明确的界定, 有很强的可比性, 强调指标体系的完整性, 并对每个指标赋予了权重, 对结果进行排序。

把科研评估与教育质量评估分开是英国评估的一个最大特色。由于主要评估科研情况, 对科研水平的把握方面相对比较准确, 评估结果不是排序, 而是对达到的水平进行定性评价。

美国学科评估是在主观评估（声誉调查）的基础上发展起来的，对声誉调查有一整套的办法，调查比较细，评估动用了比较大的财力和物力，对各个测度（指标）的相关性进行了详细的分析，有较强的事实依据。

六、开展一级学科整体水平评估的有关原则

学科水平是高等学校质量和水平的主要标志，建设世界一流大学的关键就是要建设具有世界一流水平的学科。美国加州伯克利大学前校长田长霖教授说[9]，世界上地位上升很快的学校，都是在一、两个领域首先突破。一个学校不可能在很多领域同时达到世界一流，一定要有先后，研究型大学一定要想办法扶植最最优异的学科，把它变成世界上最好的。然后其他学科也会自然而然地跟上来。加州理工学院为什么会变成这么著名大学？它的腾飞就是靠两个教授，一个是密立根，物理诺贝尔奖获得者，他使这个学校的实验物理迈上了世界一流；然后是冯·卡门，钱学森的老师，他把美国的航空技术带起来了。有这两个人开头，加上加州理工学院后来的持续发展，于是就世界知名了。

开展一级学科整体水平评估可以衡量一下我国有哪些学科能够达到世界一流水平或具有世界知名影响，更重要的是通过开展一级学科整体水平评估应当引导、促进学科发展，提升学科国际竞争力。因此，开展一级学科整体水平评估要注意把握以下几个原则。

（一）国际可比性原则

一级学科整体水平评估是在一级学科选优评估、重点学科评估的

基础上发展而来的，已经形成了自己的特色，有相当的继承性，也有相当的可操作性。就像英国、美国的学科评估一样，有自己的特色。但是，一级学科整体水平评估要有利于引导、促进学科发展，提升学科国际竞争力。因此，在探索改进一级学科整体水平评估时，既要考虑学科评估的继承性和操作性，形成自己的特色，也要放在国际环境下进行比较，指标体系应具有一定的国际可比性，同时要加强国际交流与合作，不能搞自我封闭，要请海内外专家参与评价，要积极参与国际性评估。只有这样，才能更好地引导、促进学科发展，使一批学科走向世界，成为世界知名甚至世界一流学科。

（二）指标体系优化原则

学科整体水平的评估包括的方面很多，即包括师资队伍的状况、人才培养的数量与质量、科学研究水平及成果，还包括学科本身的建设情况，学术交流、学科管理等

等，要全面评价，但更要突出重点。文献[4]指出：任何可供使用的测度（本文理解为“指标”）都有不足之处，而且不可能期望使用任何一个测度来给出有关研究生教育质量之完全满意的指标。每个尺度充其量只是围绕一个广泛概念的某个侧面的局部测度。有时，这一尺度与现实世界只有模糊和脆弱的联系，每个尺度还可能包含一大堆不相干，即与评估没有联系的“额外负担”。但通过利用若干这样的尺度以反映事物的各个不同的侧面，我们就可以限制那些不相干因素的影响并得到一个较为完整和更为逼真的事物评估结果的图像。

从系统的观点来看，指标可分为结构指标和功能指标两类[10]。结构指标反映系统内部各个组成要素的性质、状态，功能指标反映系统水平和质量的本质特征。由于系统是分层次的，每一个层次都具有相应的结构和功能，也就具有相应的结构指标和功能指标。高层次的结构指标可能是低层次的功能指标。

学科整体水平评估不同于合格评估，应突出学科的产出，突出学科的效益。对于学术性学位，应注重学术声誉和学术产出，强调评价学科的学术生产率[11]。学术生产率包含了研究生教育中研究侧面的生产力——研究生产率，以及培养研究生人才的教育侧面的生产力——教育生产率。

（三）非线性排序原则

一级学科整体水平评估结果目前是以线性顺序排列的，以得分多少排名或分出最好、次好等等。但是，以线性顺序进行的任何排列也可能是不准确的。水平大体相当的学科难于进行比较，不能简单说某个学科得分高水平就高。英国科研评估结果不是进行简单的排序，而是根据达到国际先进水平或国内先进水平的程度进行打分，这种评介更具合理性，更有利于促进学科的竞争，更有利于进行国际比较。

七、关于工科类一级学科整体水平评估的有关建议

首次开展的一级学科整体水平评估试点工作取得了成功，在广泛征求意见的基础上进一步完善，一定会取得更好的效果，更有利地引导、促进学科发展。

根据上述讨论的原则，对工科类一级学科整体水平评估提出如下建议。

（一）指标体系的进一步改进

在完善一级学科整体水平评估指标体系时应重点考虑以下几个方面：

第一，学科具有人才培养、科学研究、社会服务3种功能，一流的学科应培养一流的高素质人才，创造一流的科研成果，提供一流的社会服务。因此评价一级学科整体水平评估时也应当包括这3方面，即包含教育生产率、研究生产率和服务生产率。

评价学科的服务生产率（主要指专利和知识产权转化）主要是考虑到当前我国以经济建设为中心，学科具有直接为经济建设服务的功能，事实上很多学科主要是从事技术开发而非学术研究，把技术开发和学术研究混为一谈，不利于提升学科国际竞争力，只考虑学术研究不考虑技术开发，也将把很多学科排除在外，不利于学科发展。当然，不同学科其教育、研究和服务生产率的指标权重也会有所不同。

第二，评价学科整体水平时应主要评价学科的功能、产出、效率和效益，而不应着重评价学科的现状、条件和规模等，如教师队伍的基本状况，博士、硕士学位授权点数，重点学科数，重点实验室数等。因为，这些指标虽然非常重要，与学科整体水平有一定的相关性，但对学科整体水平的影响是间接的，而非直接的。

第三，评价学科整体水平的指标体系应具有国际可比性，不能太具有中国特色，如指标体系中不应有博士、硕士学位授权点数，重点学科数，重点实验室数等这些太具中国特色的指标，学科的划分也应尽可能与国际接轨，覆盖面不能太窄，要有利于引导学科参与国际交流、参与国际竞争。

根据上述原则，评价工科一级学科整体水平可从下面 4 个方面考虑：

- (1) 人才培养，包括培养研究生的数量和质量等；
- (2) 学术研究，包括所获基础研究和应用基础研究科研项目的数量及经费额度，发表论文质量及论文的影响，取得的创新成果等；
- (3) 社会服务，包括专利、知识产权成果转化等；
- (4) 本学科在国内外的声誉。

这 4 个方面可以作为一级学科整体水平评估的一级指标，实际上每个指标都近乎综合性评价，都可以细分为若干二级指标。

（二）关于受评学科在国内外的声誉调查

关于受评学科在国内外的声誉调查工作应从 3 方面考虑：

1. 调查的对象。

调查的对象应包括申请参加一级学科整体水平评估的有关学科的专家，国外相关学科的专家和国内企业、科研院所相关领域的专家。

2. 提供给调查对象的资料。

所有参加一级学科整体水平评估的学科应将本学科全职教学科研人员有关基本情况提供出来。英国科研评估需要提供待评估人员名单和在给定期间（某些文科为 6 年，其他学科均为 4 年）每一位科研人员的科研成果（4 篇论著或其它研究成果）；美国评估也需要列出在高等学校任职并且明显地参与了博士教育的教师名单，一般不包括退

体和辅助成员。这些数据提供给调查人员有助于调查人员的打分，也有利于防止使用兼职人员或非本学科人员的有关材料。

还应将本学科博士、硕士学位授权点数、重点实验室数、重点学科数等数据提供给调查人员。这些数据不是作为二级指标直接参与打分。

3. 有关二级指标

学术声誉一级指标下也应分若干二级指标，美国研究型博士学位点评估声誉调查分了4个二级指标。借鉴英国科研评估和美国研究型博士学位点评估，可以考虑的二级指标有：

(1) 学科教师学术水平的评估均值；

(2) 学科整体水平情况（a 大多数学科子领域为国际先进水平，其它都达到国内先进水平；b 某些学科子领域达到国际先进水平，其它均达到国内先进水平；c 大多数学科子领域达国内先进水平；d 半数学科子领域达国内先进水平；e 各学科子领域均未达国内先进水平。）

（三）关于结果排序

既可以对各个一级指标进行排序，也可以根据权重计算总的得分进行排序，权重系数可能需要进一步的完善。为解决线性排序带来的问题，可以根据声誉调查结果，对学科整体水平处于同一级并且总得分差值在一定范围内的实行并列排序。