

·工作研究·

开放实验室评议数据的分析与讨论

洪明苑

(国家自然科学基金委员会)

[摘要] 本文以数理科学部承担的国家重点实验室和开放实验室第二次运行补助费评议工作为例,对定量分和定性分的数据进行分析,给出了实验室按不同计分的排序;提出了甲、乙、丙、丁不同分档方法;讨论了定量分与定性分的相互关系;指出了当前定量评价指标和定性指标体系中的问题。此外,还反评价了各专家评议情况,不同专家评分与平均分的偏离,以及大同行与小同行在评分中带来的差异。本文对提高评价指标体系的科学性和完善评议方法进行了探讨。

1991年,国家自然科学基金委员会受国家科委委托,对国家重点实验室和开放实验室(以下简称实验室)进行了第二次运行补助费的评议工作。评议根据国家科委关于定性评议与定量评议相结合,以定性评议为主的指导原则。

在实行过程中,定量评议由“定量数据统计指标体系数据统计汇总表”^①给出。定量统计指标包括:学术水平,赋予50%权重;人才培养,赋予30%权重;开放度,赋予15%权重;大型仪器设备有效利用率,赋予5%权重。以上由各实验室自行统计给出。

定性评议则按专家评议指标体系和评议的有关原则,邀请专家评议记分投票得到。定性统计的专家评议指标体系包括:科研水平,赋予40%权重;队伍建设与人才培养,赋予30%权重;对外开放水平,赋予15%权重;科研管理水平,赋予15%权重。

综合评价分数则由定量统计分数和定性投票分数计算得出。计算时,以定性分为主,占60%;定量分占40%。根据国家科委要求,评议结果甲类实验室应占15%,难以确定资助与否的丙类实验室不超过10%,总资助率60%。

这次数理科学部承担了20个实验室的评议,邀请了31位专家,分成三个评议组:第一组为凝聚态物理的8个实验室,由11位专家组成;第二组为光学、声学、原子分子物理、核物理和核技术的8个实验室,由12位专家组成;第三组为力学和天文学的4个实验室,由8位专家组成。结果共评出甲类实验室3个,乙类9个,丙类2个,丁类6个。

由于实验室的评议问题涉及实验室的荣誉和经费的支持,是一项政策性很强的工作,而且尚缺乏经验,因此,对如何完善评议方法,提高统计指标体系的科学性,还需认真探讨。本文仅以数理科学部承担的评议工作为例,对这次评出的定量和定性分数数据进行分析与讨论。本文不与实验室评价结果直接联系,不发表原始分数,只根据需要列出一些经过处理的数据和图表。为讨论方便,实验室名称用代号 $L_{11}, \dots, L_{21}, \dots, L_{31}, \dots$;评议的专家姓名用 $P_{101}, \dots, P_{201}, \dots, P_{301}, \dots$ 代替。

① 《开放实验室运行补助费评议细则》,国家科委(91)国科发京字565号文件。

本文于1992年3月28日收到。

一、实验室评议分数的排序

首先,由于定量分是各实验室按定量数据统计指标体系和数据统计汇总表给出,因此,它属于同一系统的可比数据,由高分到低分排序即给出名次(见表1)。

表1 各实验室按不同计分排序变化的举例

| 实验室代号 | 按定量分排序 | 按综合分排序 | 按定性分排序 | 按定性分(修正)排序 | 实验室代号 |
|-----------------|--------|--------|--------|------------|-----------------|
| L ₁₁ | (1) | (1) | (1) | (1) | L ₂₁ |
| L ₁₆ | (2) | | (2) | (2) | L ₁₁ |
| L ₃₄ | | | | | L ₂₇ |
| L ₂₀ | | | | | L ₁₄ |
| L ₂₁ | (4) | (5) | | | L ₃₄ |
| L ₂₂ | | (6) | | | L ₂₀ |
| L ₁₆ | | | | (7) | L ₂₀ |
| L ₂₇ | | | (8) | | L ₂₂ |
| L ₁₇ | | | | | L ₃₃ |
| L ₂₃ | | | (10) | | L ₁₆ |
| L ₁₄ | | | | | L ₂₉ |
| L ₂₂ | | | | | L ₁₂ |
| L ₁₂ | | (13) | | (13) | L ₁₆ |
| L ₁₈ | | | | | L ₁₇ |
| L ₁₃ | | (15) | | | L ₁₉ |
| L ₃₁ | (6) | | | | L ₂₁ |
| L ₂₄ | | | (17) | | L ₂₄ |
| L ₂₅ | (8) | | | | L ₁₈ |
| L ₂₆ | (19) | | (19) | (19) | L ₂₆ |
| L ₂₃ | | (20) | | | L ₂₂ |

—— 变化不大; ---- 急剧下降; - - - - 急剧上升

第二,关于定性分的排序,由于把20个实验室分成三组进行评议,严格说,各组标准掌握是有所差别的,直接排序是否合适尚有问题。为此,本文试图用三组总平均分与分组平均分数差别加以修正,即第一组平均分偏高1.31分,对第一组各实验室均扣除1.31分;第二组偏低3.18分,相应给各个实验室增加此分数;第三组偏高1.87分,也相应予以扣除。以上所得结果为定性分(修正)。为了相互比较,表1给出了各实验室按不同计分排序的变化。看来,这对于实验室评价有一定参考价值。

第三,关于甲、乙、丙、丁的分档,可用相对分数,也可用绝对分数。国家科委要求按相对分数分档,即对被评议实验室按分数排序后,划出甲类15%,乙类45%,丙类10%,丁类

30%。我们评议时考虑到分组的因素，采用相对分数排序和投票的方式，确定出3个甲类、2个丙类、9个乙类和6个丁类实验室。这种相对分数确定档次的办法会发生这样的情况：即使实验室都很好，也会有丁类，反之实验室都很差，也有甲类。为避免上述情况，笔者认为可用绝对分数分档，比如，规定高于90—100分为甲类（特优）；高于80—90分为乙类（优）；高于70—80分为丙类（良）；高于60—70为丁类（中）；低于60分为戊类（差）。表2给出四种不同方法分档的结果。

表2 不同分档方法的结果

| 分 档 方 法 次 | 实验室数目 及百分比 | 按 相 对 分 数 分 档 | | 按 绝 对 分 数 分 档 | |
|--------------|---------------|---------------|------------------|---------------|------------------|
| | | 按 综 合 分 | 按 综 合 分 加 投 票 | 按 综 合 分 | 按 定 性 分 (修 正) |
| 甲 | | 3 | 3 | 1 | 3 |
| | | 15% | 15% | 5% | 15% |
| 乙 | | 9 | 9 | 7 | 11 |
| | | 45% | 45% | 37% | 55% |
| 丙 | | 2 | 2 | 10 | 5 |
| | | 10% | 10% | 53% | 25% |
| 丁 | | 6 | 6 | 1 | 1 |
| | | 30% | 30% | 5% | 5% |

由表2可以看出，不同分档方法结果是不同的。似乎以绝对分数分档（无论用综合分或定性加分修正分）可能更接近于实际，即“特优”和“中”是少数，大多数实验室属于优和良。这也符合专家们关于扩大补助费授予面的要求。

二、定量分与定性分的关系

在上节中已经看到被评议的实验室按分数排序的情况，这里进一步用图表加以分析。

图1表示每一组实验室排序情况。分布在45°斜线上的有 $L_{11}(1,1)$, $L_{16}(3,3)$ 。这表明，这两个实验室在该组内定量分排序与定性分排序是一致的；斜线左上方的 $L_{15}(2,4)$ 和 $L_{17}(4,6)$ 实验室，其定量分名次在前，定性分名次在后；斜线右下方的 $L_{14}(5,2)$ 和 $L_{13}(8,5)$ ，则定量分名次在后，定性分名次在前。

现将20个实验室总排序画在同一图上(图2)。可以看出，约60%的实验室分布在接近于45°斜线附近，从而可以认为，定量分排序和定性分排序具有一定相关性，两者结合考虑，作为划分实验室评价的档次，基本上是可行的；而 L_{15} , L_{21} , L_{32} , L_{31} , L_{26} , L_{33} 6个实验室，占总数30%，偏离45°斜线较远，这表明两种分数排序相互矛盾较大，有必要对统计指标体系进一步深入剖析，探索引起这种矛盾的原因，以便改进统计指标体系。

下面进一步按学术水平、研究队伍、开放度、仪器利用率及管理水平逐项分析，并用图解表示在图3, 4, 5中。图左边纵坐标高度表示分项分数，右边纵坐标高度表示总分数，顶端实线表示定量分，虚线表示定性分，“S”表示学术水平，“D”表示研究队伍，“K”表示开放度，“Ga”表示仪器利用率，“Gb”表示管理水平，“Z”表示总分。现把图2中接近斜线的，也就是定量分排序与定性分排序接近的 $L_{11}(1,2)$ 和 $L_{25}(18,19)$ ，表示在图3上，而斜线左上方的

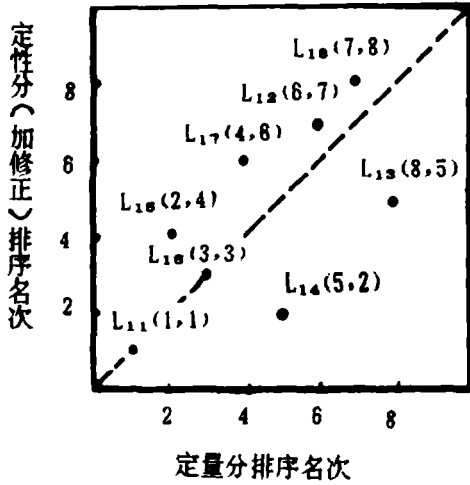


图1 第1组实验室名次排序

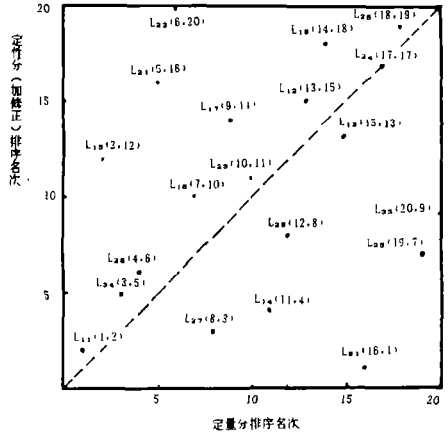


图2 三组实验室名次总排序

$L_{21}(5,16)$ 表示在图4上;斜线右下方的 $L_{31}(16,1)$ 表示在图5上。

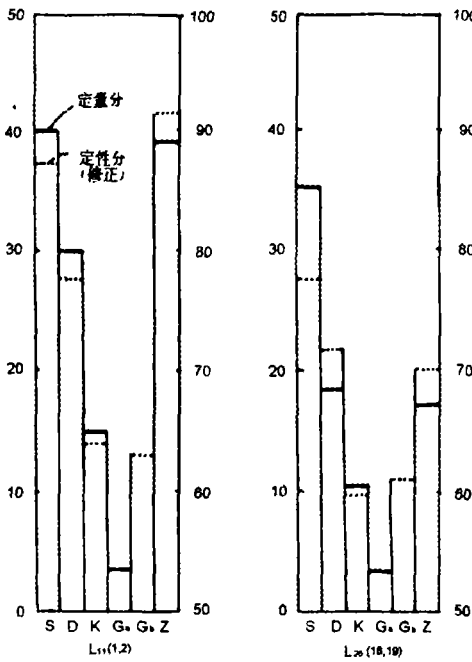


图3 $L_{11}(1,2)$, $L_{25}(18,19)$ 实验室分项分数图解

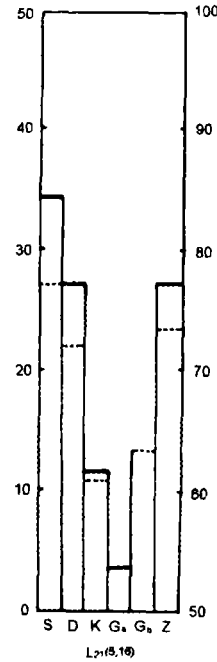


图4 实验室 $L_{21}(5,16)$ 分项分数图解

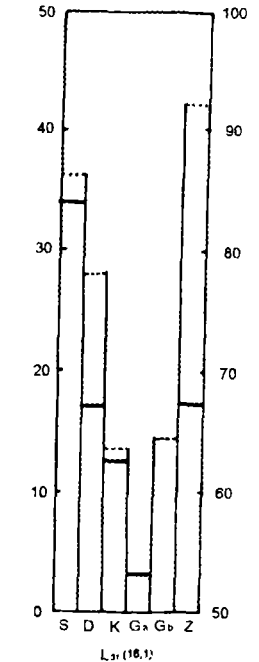


图5 实验室 $L_{31}(16,1)$ 分项分数图解

图3中 $L_{11}(1,2)$ 实验室的定量分与定性分比较接近; $L_{25}(18,19)$ 定性分中, 学术水平分下降与研究队伍分提高, 总分变化不大, 其排序不发生大的变动。图4中的 $L_{21}(5,16)$ 由于定性分各项全面下降, 造成总排序名次大大后移。图5中 $L_{31}(16,1)$ 的总的定性分远远高于定量分, 从而在总排序上名次大幅度上升。

从图3—5可能引出这样的看法，实验室办得好坏，关键在于学术水平以及与之相关的研究队伍，这也是定性评价与定量评价差别最大的地方。定量分是按论文和成果的数量给出分数，比较容易进行统计。但论文和成果的水平就难以给出合适的分值，即使考虑了一定权重，但实质上存在评价的“非线性效应”。比如一篇国际上获得领先地位的创新性论文，其评分则不是简单等于几篇或几十篇论文分数之和。因而必须考虑学术水平和研究队伍水平评价的“非线性因素”，对目前两个指标体系做适当修改，使其更为一致。

至于开放度和管理水平，各实验室有一定差别，评价变化幅度不大；仪器利用率一项由于只赋予5%权重，差别也就很小，实际上对评议几乎不起调节作用。

三、对专家评议情况的分析

表3分别给出3个评议组专家对各个实验室给分与全组专家给分平均值的偏离情况。从最大偏离一栏中，可以看出某专家对该组实验室评分偏高或偏低的最大值。把偏离绝对值平均，一定程度上可以反映各专家判断的准确性。总的，大约一半专家评分比较适中，相当一部分专家偏离幅度较大。当然，在客观上与这次评议会时间较短，看材料、讨论不甚充分有关，但也与专家的判断力有关。如果我们经常进行这种分析，将有助于今后对实验室评议专家的选择。

表3 三组专家对实验室记分与平均分的偏离

| 第一组 专家 | 最大偏离 | | 偏离绝对 对平均值 | 判断 | 第二组 专家 | 最大偏离 | | 偏离绝对 平均值 | 判断 | 第三组 专家 | 最大偏离 | | 偏离绝对 平均值 | 判断 |
|------------------|------|-------|--------------|-----|------------------|-------|-------|-------------|-----|------------------|-------|------|-------------|-----|
| | + | - | | | | + | - | | | | + | - | | |
| P ₁₀₁ | +7.4 | -0.1 | 3.0 | 偏高 | P ₂₀₁ | +8.8 | -0.3 | 3.4 | 稍高 | P ₃₀₁ | +6.5 | -4.5 | 4.1 | 稍高 |
| P ₁₀₂ | | -14.3 | 5.5 | 偏低 | P ₂₀₂ | +10.2 | -14.2 | 7.2 | 偏高 | P ₃₀₂ | +0.6 | -5.3 | 3.7 | 较适中 |
| P ₁₀₃ | +0.9 | -9.6 | 2.8 | 较适中 | P ₂₀₃ | +3.4 | -6.4 | 2.4 | 适中 | P ₃₀₃ | +13.1 | -0.2 | 6.4 | 偏高 |
| P ₁₀₄ | +2.3 | -14.4 | 3.5 | 偏低 | P ₂₀₄ | +11.2 | -3.0 | 5.3 | 偏高 | P ₃₀₄ | | -9.7 | 6.0 | 偏低 |
| P ₁₀₅ | +2.9 | -4.9 | 2.2 | 适中 | P ₂₀₅ | +4.3 | -4.6 | 2.7 | 适中 | P ₃₀₅ | +8.3 | | 4.7 | 稍高 |
| P ₁₀₆ | +8.7 | | 4.9 | 偏高 | P ₂₀₆ | +2.2 | -12.9 | 4.2 | 偏低 | P ₃₀₆ | | -9.0 | 5.7 | 偏低 |
| P ₁₀₇ | +4.5 | | 2.8 | 较适中 | P ₂₀₇ | +1.7 | -8.3 | 3.1 | 较适中 | P ₃₀₇ | +7.9 | -0.6 | 3.0 | 较适中 |
| P ₁₀₈ | +7.8 | -9.9 | 4.7 | 偏高 | P ₂₀₈ | +5.2 | -4.5 | 3.2 | 较适中 | P ₃₀₈ | +3.3 | -6.9 | 3.3 | 较适中 |
| P ₁₀₉ | +3.0 | -1.7 | 1.6 | 适中 | P ₂₀₉ | +2.5 | -7.8 | 3.3 | 较适中 | | | | | |
| P ₁₁₀ | +2.4 | -4.3 | 1.7 | 适中 | P ₂₁₀ | +5.1 | -4.6 | 3.5 | 稍低 | | | | | |
| P ₁₁₁ | +6.7 | -4.4 | 3.7 | 偏高 | P ₂₁₁ | +2.9 | -18.3 | 6.8 | 偏低 | | | | | |
| | | | | | P ₂₁₂ | +5.0 | -9.4 | 4.1 | 稍低 | | | | | |

由于被评议的实验室在学科内容上跨度较大，还有必要考查一下专业上的差别对评分的影响。除第一组属凝聚态物理外，第二组和第三组均存在大同行和小同行问题。表4列出了小同行专家给出平均分与全组专家平均分的偏离情况。可以看出，除L₂₆实验室的4位小同行给出分数比该组平均分低1.9分外，其余11个实验室的小同行均给出了比其所在组平均分略高的分数。因此，专业不同，在一起评议时有可能使实验室得分较低。另外，将第二、三组与第一组相比，专业不同的专家给出的分数有着分散度较大的趋势。这些在今后组织实验室评议时也要加以注意。

表4 第二、三组小同行专家平均分与全组平均分的偏离

| | 第二组 | | | | | | | | 第三组 | | | |
|-------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | L ₂₁ | L ₂₂ | L ₂₃ | L ₂₄ | L ₂₅ | L ₂₆ | L ₂₇ | L ₂₈ | L ₃₁ | L ₃₂ | L ₃₃ | L ₃₄ |
| 小同行人数 | 6 | 5 | 5 | 5 | 7 | 4 | 3 | 3 | 4 | 4 | 4 | 4 |
| 分数偏离 | +2.0 | +2.2 | +2.3 | +4.7 | +1.4 | -1.9 | +2.2 | +3.6 | +1.4 | +1.1 | +0.6 | +0.5 |

以上讨论, 仅作为尝试, 以期抛砖引玉, 引起更多人的关注, 共同讨论, 改进工作。

ANALYSIS AND DISCUSSION ON THE DATA FROM THE APPRAISALS OF OPEN LABORATORIES

Hong Mingyuan

(National Natural Science Foundation of China)

Abstract

This article analyses and discusses the data from the appraisals of Key Laboratories and Open Laboratories. It gives the order of the laboratories by marks, and also their grade in A, B, C or D. It discusses the relation between quantitative marks and qualitative marks and points out some problems in the quantitative index system and the qualitative index system. Moreover, it discusses the general features of the appraisals, the deviations of their marks from the average, and the difference of marks between the group from the small research field and that from the large research field. Then, it probes into an improvement of the scientific evaluation index system.

·信息·

日本发表1992年度科学技术研究调查结果

日本政府自1983年以来, 每个财政年度(每年4月1日至下一年3月31日)对全国的科学技术研究的实际情况进行调查。

1992年全国研究经费总额为13兆7715亿日元(折合1187亿美元), 比上一年度增加5.3%, 研究经费占国民生产总值(GNP)已连续两年达到3%。

各研究所的研究经费增加7.0%, 达到139.7亿美元。如果按学科划分, 工程科学的研究经费占48%, 为66.6亿美元。研究所的研究经费使用比例为: 基础研究费16.1%; 应用研究费28.3%; 开发研究费55.6%。各研究所专职科研人员人均研究经费比上年度增加4.1%, 平均达到32.3万美元。其中理科人均研究经费最高达到55.3万美元。

高等学校1992年度研究经费为2兆4079亿日元(207.6亿美元), 比上年度增加4.8%。研究经费使用比例为: 基础研究费53.8%; 应用研究费38.4%; 开发研究费8.8%。如果按学科划分: 理科方面的基础研究费86.6%; 应用研究费9.4%; 开发研究费4.0%。高等学校专职人员人均研究经费为9.7万美元, 比上年度增加2.7%。理科方面研究经费最高为13.8万美元。

(祖广安 编译)