

·基金纵横·

国家自然科学基金同行评议结果评价方法与专家遴选因素分析

王国彪 彭芳瑜

(国家自然科学基金委员会工程与材料科学部, 北京 100085)

2008年国家财政预算计划投入自然科学基金53.59亿元,比上年增长约25%^[1]。在国家对基础研究重视程度与日俱增、投入不断加大的背景下,如何管好用好科学基金,成为大家关注的焦点问题。为此,国家自然科学基金委员会陈宜瑜主任在六届一次全委会的工作报告中提出:将适时启动“科学基金资助管理绩效国际评估”工作^[2]。同行专家评议(peer review)尽管存在一些缺陷^[3,4],但仍然是遴选竞争类项目的一种有效方式,是基金项目评审工作的重要环节。因此,如何公正、合理地遴选同行评议专家,确保评议结果的权威性与有效性,显得非常重要。文献[5]采用专家的基本情况、修养和工作业绩作为选择同行评议专家的依据,文献[6]则从统计学与集合论的角度给出了同行评议专家定量评估指标。然而,上述研究工作侧重于对同行评议专家个体进行分析、评估,无法对同行评议专家群体(某一类同行评议专家)进行整体评估。

本文通过定义三个定量评价指标(项目熟悉度、项目资助度、项目综合评价指标偏差),以同行评议专家的职称、学位、年龄、当年是否申报项目作为主要影响因素,以2008年度机械学科基金项目的数据为样本,定量分析了影响因素对评价指标的影响度,以期建立同行评议结果的评价方法和专家遴选的基本原则,为科学基金资助管理绩效评估提供参考。

1 同行评议专家评议结果的评价指标

1.1 同行评议专家评议结果的定量评价指标

为反映同行评议专家评议结果的可信度(Credibility)、公正性(Justifiability)、以及有效性(Effectiveness),本文对应定义三个定量评价指标,即项目熟悉度、项目资助度、项目综合评价指标偏差,以便定量计算与分析同行评议专家的评议结果。

(1) 项目熟悉度:假设有 m 位同行评议专家,第 i 位专家的熟悉率 cd_i 和这 m 位专家的平均熟悉率 CD 分别定义为:

$$cd_i = \frac{n_{ci}}{n_i} \times 100\%, \quad CD = \frac{\sum_{i=1}^m cd_i}{m} \quad (1)$$

式中, $n_i, n_{ci} (i=1, 2, \dots, m)$ 分别为第 i 位专家的项目指派数量和认为熟悉的项目数量。与此类似,可以定义这 m 位专家的平均较熟悉率和平均不熟悉率。这三个指标构成这 m 位专家对项目熟悉程度的评价指标,一定程度上反映评议结果的可信度。

(2) 项目资助度:第 i 位专家的优先资助率 jt_i 和这 m 位专家的平均优先资助率 JT 可分别定义为:

$$jt_i = \frac{n_{ji}}{n_i} \times 100\%, \quad JT = \frac{\sum_{i=1}^m jt_i}{m} \quad (2)$$

式中, $n_{ji} (i=1, 2, \dots, m)$ 为第 i 位专家判定优先资助的项目数。与此类似,可以定义这 m 位专家的平均可资助率和平均不予资助率。这三个指标构成这 m 位专家对项目资助程度的评价指标,一定程度上反映评议结果的公正性。

(3) 项目综合评价指标偏差:第 i 位专家的平均综合评价指标偏差 et_i 和这 m 位专家的平均综合评价指标偏差 ET 可分别定义为:

$$et_i = \frac{\sum_{j=0}^{j=n_i} |v_{ji} - \bar{v}_j|}{n_i}, \quad ET = \frac{\sum_{i=0}^{i=m} et_i}{m} \quad (3)$$

式中, $\bar{v}_{ji} (j=1, 2, \dots, n_i; i=1, 2, \dots, m)$ 为第 i 位专家对指派的第 j 个项目的综合评价指标(综合评价指标中的优、良、中、差,对应分值分别为4, 3, 2, 1), $\bar{v}_j (j=1, 2, \dots, n_i)$ 为第 j 个项目所有专家综合评价指标的平均分。平均综合评价指标偏差可以反

本文于2008年8月6日收到。

映这 m 位专家与其他专家相比,对项目综合评价指标的差异,一定程度上反映评议结果的有效性。

2008 年度机械学科共遴选了 1 194 位同行评议专家,指派评议项目 11 405 份,按期返回 11 247 份评议结果。其中,对评议项目“熟悉”的有 5 993 份,“较熟悉”的有 4 462 份,“不熟悉”的有 792 份,由式(1)可得,平均熟悉率(CD)为 53.3%,平均较熟悉率为 39.7%,平均不熟悉率为 7.0%;对评议项目给出“优先资助”的有 1 421 份,“可资助”的有 3 525 份,“不予资助”的有 6 301 份,由式(2)可得,平均优先资助率(JT)为 12.6%,平均可资助率为 31.3%,平均不予资助率为 56.0%;由式(3),可以计算平均评价偏差(ET)为 0.523。这些值可作为评价某类同行评议专家的标准指标,如果某类同行评议专家的评价指标与相应的标准指标相差过大,例如,年龄小于 35 岁的同行评议专家的平均不熟悉率为 12.5%,与同行评议专家的平均不熟悉率 7.0% 相差过大,可以认为小于 35 岁的同行评议专家在项目熟悉度上不是很好,若想提高这个指标,就应慎重选择 35 岁以下的专家。

1.2 同行评议专家遴选因素对评价指标的影响

同行评议专家评议结果受专家的职称、学位、年龄等诸多因素影响。为了进一步分析这些因素对评议结果评价指标的影响程度,本文提出了变化率(Variation)、波动值(Fluctuation)、敏感度(Sensitivity)三个定量指标。在此以平均熟悉率为例,说明这三个指标的计算方法。

若专家的平均熟悉率为 CD ,按照某一影响因素(如同行评议专家的职称)将同行评议专家分为 N 组,第 k 组同行评议专家的平均熟悉率为 CD_k ($k = 1, 2, \dots, N$),那么第 k 组同行评议专家平均熟悉率的变化率 VT_k 、这 N 组同行评议专家平均熟悉率的波动值 FT 分别为:

$$\begin{cases} VT_k = \frac{CD_k - CD}{CD} \times 100\% & (k = 1, 2, \dots, N) \\ FT = \max\{VT_k, k = 1, 2, \dots, N\} \\ \quad - \min\{VT_k, k = 1, 2, \dots, N\} \end{cases} \quad (4)$$

将不同影响因素(如同行评议专家的职称、学位、年龄等)对应的平均熟悉率波动值从小到大进行排序,并分别赋予其 1, 2, 3, ..., 该值即为影响因素对平均熟悉率的敏感度,该值越大说明影响越大,在遴选同行评议专家时就要重点考虑这个因素。

2 同行评议专家的分布状况及其对评价指标的影响

2.1 同行评议专家的职称分布及其对评价指标的影响

在 2008 年度机械学科遴选的 1194 位同行评议专家中,正高职称 871 人,占 72.9%,副高职称 316 人,占 26.5%,中级及其他职称 7 人,占 0.6%。同行评议专家主要以副高及以上职称的为主,中级及其他职称的同行评议专家人数很少,分析时将不予考虑。

如图 1(a, b)所示,正高职称的平均熟悉率为 55.5%,平均较熟悉率为 38.4%;副高职称的平均熟悉率为 47.1%,平均较熟悉率为 43.1%,两项合计比正高职称的低 3.7%。正高职称的平均优先资助率较高,为 13.5%,副高职称的则较低,为 10.3%;副高职称的平均不予资助率较高,为 57.9%,正高职称的较低,为 55.4%。正高职称的平均综合评价指标偏差较高,为 0.533,副高职称的则较低,为 0.496。

由此可知,正高职称同行评议专家评议结果的可信度相对较高,对项目的资助度相对较松。而同行评议专家的职称对评议结果的有效性影响不大。

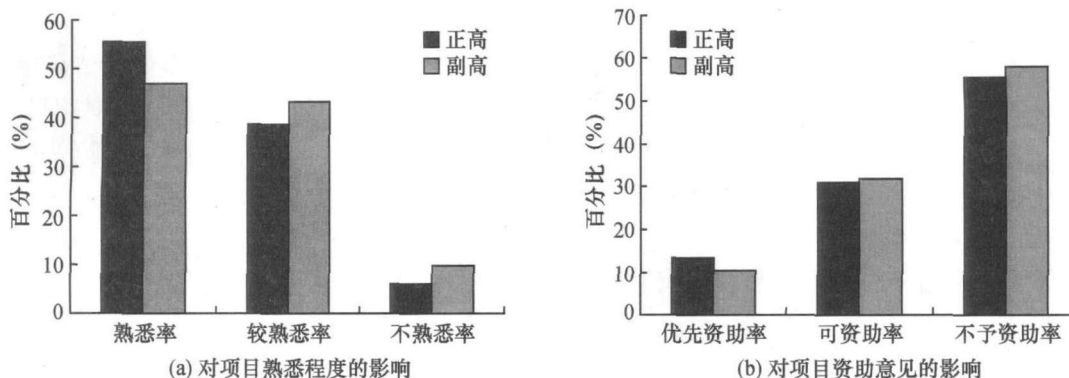


图 1 同行评议专家的职称分布及其对项目评价指标的影响

2.2 同行评议专家的学位分布及其对评价指标的影响

2008年度机械学科同行评议专家以博士为主,共1041人,占87.2%,硕士87人,占7.3%,学士及其他66人,占5.5%,后两者合并一起,暂称为其他学位。

如图2(a, b)所示,博士学位的平均熟悉率为53.7%,平均较熟悉率为39.3%,合计为93.0%,仅比其他学位的高0.6%,说明学位对项目的平均熟

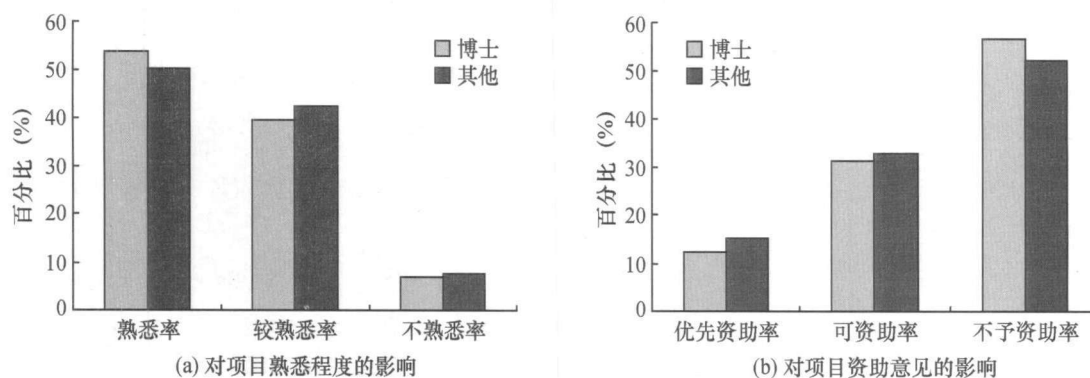


图2 同行评议专家的学位分布及其对项目评价指标的影响

2.3 同行评议专家的年龄分布及其对评价指标的影响

2008年度机械学科同行评议专家主要是以55岁以下的为主,其年龄结构分布如图3(a)所示。如图3(b, d)所示,35岁以下的同行评议专家对评议项目的不熟悉率最高,为12.5%,熟悉率最低,为43.7%;不熟悉率最低的主要分布在41~50岁、55~60岁、以及66~70岁之间,分别为5.7%、4.0%、以及0.7%;平均优先资助率较高的主要分布在46~50岁的,为15.5%,61岁以上的平均优先资助率也较高,而较低的则分布在56~60岁,为9.1%;平均不予资助率的平均值基本上随着年龄的增加呈下降趋势,35岁以下的最高,为59.5%,70岁以上的最低,为50.0%;平均综合评价指标偏差较高的主要分布在66~70岁和70岁以上的,分别为0.557、0.577,51~55岁以上的也较高,而较低的则分布在30~35岁,为0.484。

由此可知,同行评议专家的年龄对评议结果评价指标的影响相对都较大。同行评议专家所处的年龄段不同,对项目的熟悉度、资助度以及综合评价指标偏差都呈现较强的波动性。

2.4 同行评议专家当年是否申报项目情况及其对评价指标的影响

当年申报项目且申报代码与被评议的项目申请

代码相同的专家有264人,占22.5%;当年申报项目但申报代码与被评议的项目申请代码不同的专家有146人,占12.4%;当年没有申报项目的专家有764人,占65.1%。可以看出当年既是基金项目的申报者又是评议者的专家占了一定的比例。

由此可知,具有博士学位的同行评议专家对项目的资助度相对较紧,而学位对项目评议结果的可信度和有效性均影响不大。

如图4(a)所示,当年申报且代码相同的对评议项目的平均熟悉率较高(59.9%),申报但代码不同的对评议项目的平均熟悉率较低(48.5%),而平均不熟悉率较高的也是申报但代码不同的,为9.4%。如图4(b)所示在优先资助率方面,没有申报项目的最高(13.4%),其次是申报且代码相同的(11.9%),申报但代码不同的最低(10.2%)。当年是否申报项目对平均综合评价指标偏差的影响不大,申报但代码不同的略高(0.527),没有申报项目的同行评议专家略低(0.521)。

由此可知,当年申报项目但申请代码与被评议项目代码不同的同行评议专家,在对项目的熟悉度、平均综合评价指标偏差等指标上都较差。因此,为保证评议结果的合理性,要求尽可能选择小同行专家(即专家库中专家的学科代码与被评议项目的申请代码相一致)。

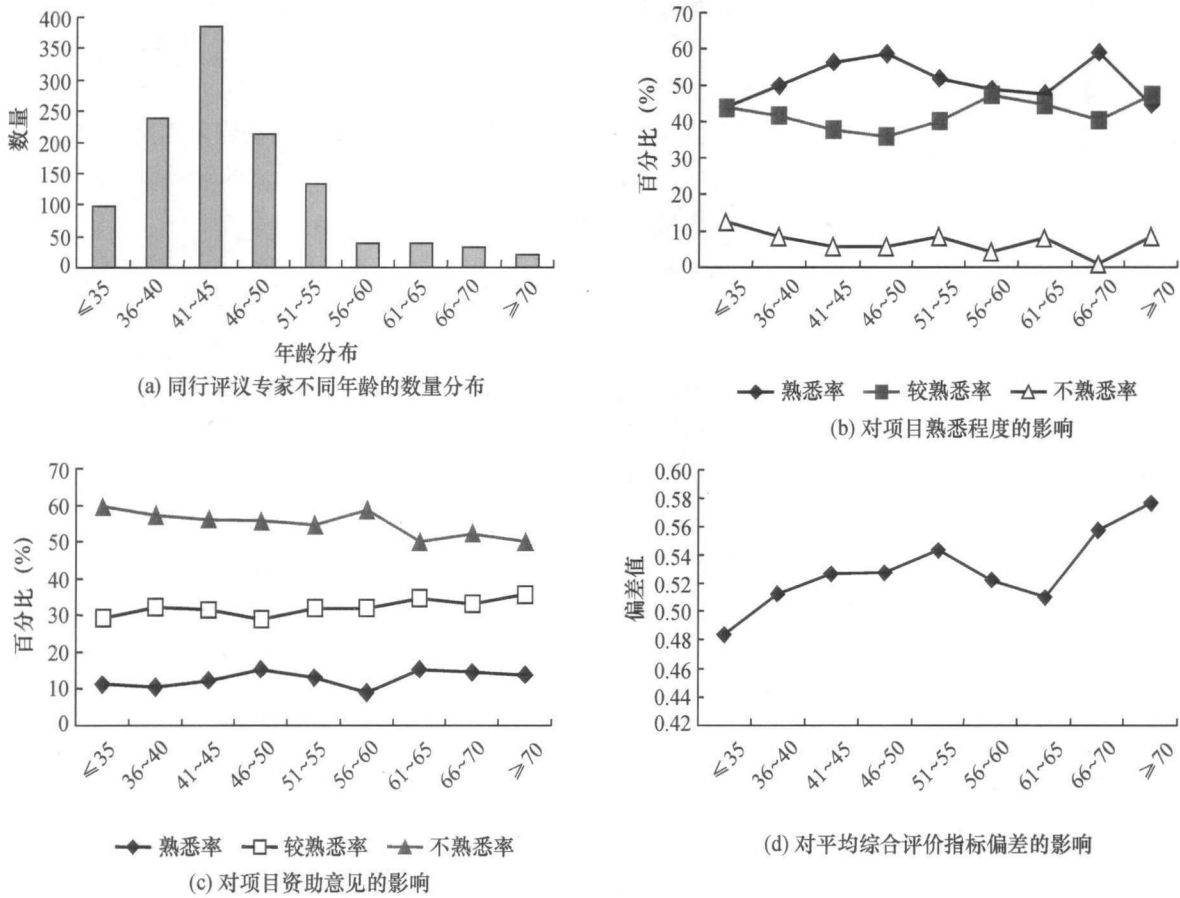


图3 同行评议专家的年龄分布及其对项目评价指标的影响

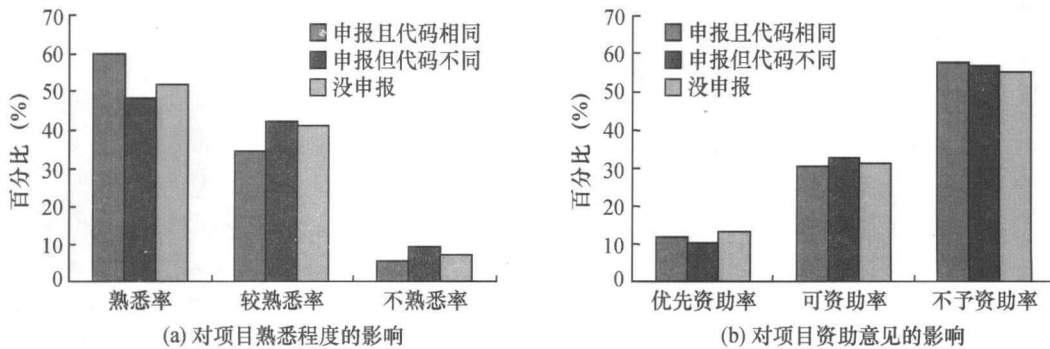


图4 同行评议专家当年是否申报项目对项目评价指标的影响

3 同行评议专家遴选因素对评价指标敏感度的分析

3.1 同行评议专家遴选因素对评价指标的波动值分析

通过式(4)可计算各因素对评价指标的变化率和波动值,以便进一步分析同行评议专家遴选因素对评价指标的影响。以年龄对平均不熟悉率的影响为例,35岁以下的平均不熟悉率最高,为12.45%,66—70岁的平均不熟悉率最低,为0.67%。以所有同行评议专家的平均不熟悉率7.04%为标准值,35岁以下的平均不熟悉率的变化率为76.76%,66—70岁的平均不熟悉率的变化率为-90.50%,因此,同行评议专家的年龄对平均不熟悉率的波动值为

111.13%。依次类推,可以获得各影响因素对各评价指标的波动值,如图5所示。

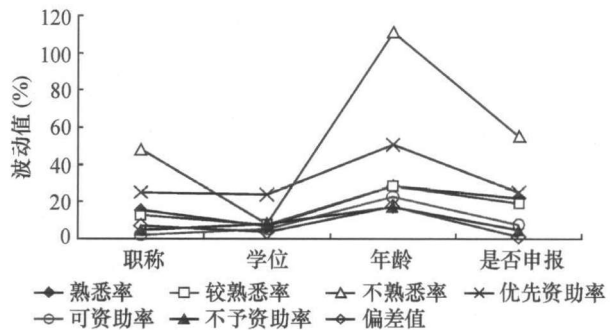


图5 熟悉度、资助度、以及综合评价指标偏差等评价指标对各影响因素的波动值

3.2 同行评议专家遴选因素对评价指标的敏感度分析

根据图5的波动值,可进一步得出各因素对评价指标的敏感度。以平均不熟悉率为例,在本文考虑的4个因素中,年龄对平均不熟悉率的波动值最大,为111.13%,故年龄对平均不熟悉率的敏感度

为4,学位对平均不熟悉率的波动值最小,为7.79%,故学位对平均不熟悉率的敏感度为1。同理,当年是否申报项目、职称的敏感度分别为3、2。依此类推,可以得到各影响因素对各评价指标的敏感度,如表1所示。由此可知,年龄对评价指标的敏感度影响最大。

表1 各影响因素对评价指标的敏感度

影响因素	项目熟悉度			项目资助度			平均综合评价 指标偏差
	平均熟悉率	平均较熟悉率	平均不熟悉率	平均优先资助率	平均可资助率	平均不予资助率	
职称	2	2	2	2	1	2	3
学位	1	1	1	1	2	3	2
年龄	4	4	4	4	4	4	4
当年是否申报项目	3	3	3	3	3	1	1

3.3 2008年度机械学科同行评议专家遴选情况分析

(1) 同行评议专家的年龄对所评议项目的熟悉度影响最大,年龄小于35岁的和年龄大于70岁的对所评议的项目的不熟悉程度相对较高,其次是同行评议专家当年是否申报项目,主要是当年申报项目但代码与所评议项目的申请代码不同的专家对所评议项目的不熟悉程度较大。

(2) 同行评议专家对所评议项目资助意见的资助度而言,专家的年龄对这个指标影响最大,年龄大于45岁的优先资助率相对较高;其次是同行评议专家当年是否申报项目,没有申报项目的同行评议专家的优先资助率较高,而当年申报了项目的优先资助率则相对较低。

(3) 同行评议专家对所评议项目评价的有效性而言,同行评议专家的年龄对这个指标影响最大,具体表现为年龄超过66岁的同行评议专家对项目评价的有效性较低;其次是职称的影响。

由表1可以进一步看出,年龄对评议结果各个指标的敏感度都为4,说明普遍影响较大,其次是当年是否申报项目和职称,而学位的影响普遍较小。也即同行评议专家的年龄、当年是否申报项目对评议结果的可信度、公正性、以及有效性有较大影响。

4 结束语

(1) 同行评议专家评议项目数量过多,规定在一个月内完成所有项目的评审,评议专家工作压力较大,一些专家仅对其感兴趣的项目进行认真评议,对不太感兴趣的项目,评议往往流于形式,也就失去

了对项目客观评价的意义。

(2) 对一些研究面宽广、研究内容发展很快,或者是专业性不是很强的领域,如仿生设计、微纳设计与制造、检测技术、非传统加工工艺与装备等,研究涉及面广泛,原理千差万别,使专家的选择、以及专家对项目创新性的评议产生一定的困难,专家评议意见差别较明显。

(3) 尽快建立同行评议专家评议意见的综合评价体系,以便于学科对遴选的同行评议专家进行评价,以利于今后同行评议专家的遴选。信息中心应增加对同行评议专家评议意见的分类检索功能,以便于学科进行相关的分析统计。

(4) 学科应定期分析、总结同行评议专家评议意见,及时更新那些不认真评议的专家,保持学科同行评议专家库的时效性和权威性。

本文仅以一年的同行评议数据为样本,提出的分析方法和得出的结论有待于今后工作的进一步检验,今后仍需继续观察、分析,以便找出具有普适性的规律,为科学基金资助管理与绩效评估提供参考。

参 考 文 献

- [1] 陈宜瑜. 坚持解放思想,推动科学发展,为提升国家科技竞争力作出更大贡献——在六届一次全委会上的工作报告. 2008年5月27日.
- [2] 陈晨. 以稳定资助保证学科均衡发展——陈宜瑜主任谈新时期基金工作. 科学时报, 2008年6月16日.
- [3] 刘明. 同行评议刍议. 科学学研究, 2003(6): 574—580.
- [4] 常健民. “同行专家评议”的举措是科技创新的“紧箍咒”. <http://blog.sina.com.cn/s/blog-4d1da4af010082wg.html>.
- [5] 王志强. 关于完善同行评议制度的若干问题和思考. 中国科学基金, 2002, 16(5): 309—313.
- [6] 王成红, 何杰, 刘克, 宋苏. 关于同行评议专家定量评估指标研究的几个新结果. 系统工程理论与实践, 2004, 2: 83—89.

ANALYSIS OF EVALUATING METHOD AND EXPERTS SELECTED FACTORS OF PEER REVIEW FOR PROPOSALS OF NSFC

Wang Guobiao Peng Fangyu

(Mechanical Engineering Division, National Natural Science Fund of China, Beijing 100085)