

文章编号: 2095-1663(2018)04-0041-05

公平公正视角下的研究生招生面试问题研究

唐 润

(南京财经大学 管理科学与工程学院, 南京 210023)

摘要: 面试是研究生招生过程中选拔人才的一个重要环节,如何保证面试结果的公平公正性是一个非常值得研究的问题。本文首先分析影响面试考官评价公平公正性的主要因素,并在此基础上构建研究生面试评价指标体系,接着提出一种基于专家评价意见贴近度动态调整专家权重的方法,最后通过案例对比分析应用该算法对评价结果的修正作用,验证了该算法可以在一定程度上减少某些专家在对面考生进行评分时可能存在的宽大效应和严格效应。

关键词: 公平公正;研究生招生;权重;面试

中图分类号: G643 **文献标识码:** A

近年来,我国的考研人数不断增加。据统计,2016年全国研究生考试的报名人数达到177万人,2017年则增加到201万人,而研究生录取人数基本维持在51万左右,2018年全国硕士研究生招生考试的报考人数238万人,比2017年增加了37万人,因此考研竞争变得更加激烈。如何按照教育部的要求,强化复试综合考核,注重对考生科研创新能力、综合素质和一贯表现的考查,严明招生纪律,加强监督检查,确保招生录取规范有序、公平公正,是一项非常重要的任务^[1]。面试作为研究生复试工作中非常重要的一个环节,可以直接考察考生除了应试能力之外的多个方面,但面试评价的主观性也非常大,如何建立科学的评价体系和选择合理的评价方法,从而避免在面试环节中存在某些影响公平公正现象的发生非常值得研究。

一、影响面试考官评分公平公正性的因素分析

在研究生面试环节中,一般采用专家小组的方式进行面试。面试组专家是研究生招生面试中的主导者,考生是被动的参与方和被选择对象,各个专家

对考生的评分高低会直接影响到该考生能否被录取,一旦专家组成员的评分结果一致性低,必然存在两个效应:宽大效应和严格效应^[2-3]。宽大效应的形成原因是某位专家对某个考生的评价高于其实际表现,分数明显高于组内其他专家,最后在分数加总的时候导致该考生获得的实际分数要高于其他考生,造成了对其他考生的不公平结果。严格效应是指某专家对某个考生的评价低于其实际表现,评价分数明显低于组内其他专家,导致该考生的分数加总结果偏低,这对于该考生来说同样也是不公正的。

宽大效应和严格效应大多是因为专家个人的主观臆断造成的结果,其背后存在多种原因,比如有些面试专家会将自己的个性特点投射到面试过程中,遇到和其个性非常类似的考生可能无形当中因为第一印象较好而给出更高的分数,弱化了考生在面试全过程中的表现;也有一些专家会因为考生某个方面比较突出忽略了考生的整体素质,从而做出对考生综合素质不匹配的评价;另外,在中国这样一个人情社会的大背景下,也存在个别考生家长利用社会关系给某个面试专家事前打招呼的现象,最后获得了该专家的评价高分。因此专家评价的一致性对于

收稿日期:2018-04-25

作者简介:唐润(1979—),男,安徽桐城人,南京财经大学管理科学与工程学院副院长,副教授,博士。

基金项目:南京财经大学教学改革项目(编号:Y1212,T_RXW110011);南京财经大学学位与研究生教育重点项目(编号:Y180004)

解决面试过程中的公平公正问题非常重要。

二、考研面试评价指标体系构建

在面试的过程中,如果不设立系统的评价指标体系,仅仅由面试专家按照对考生整体表现的主观印象进行总体评分,面试过程中的宽大效应和严格效应会表现得更加明显,因此有必要建立科学合理的面面试评价指标体系。

(一) 指标体系设计的原则

1. 系统性原则

研究生招生面试是由专家组直接面对面地考核考生的综合素质,考生素质涉及到研究生后期培养过程中的诸多问题,其评价指标应该以为国家选拔出优秀人才作为整体目标来进行设计,从这一整体目标出发,遵循系统性原则,将整体目标逐层分解若干子要求,转化成评价指标后形成系统结构,这个系统结构必须层次分明、覆盖全面,结构分支之间的逻辑关系清晰。

2. 可行性和适度性原则

可行性指的是在设计评价指标时要做到尽量简单实用,易于操作,考核要素的内涵具体明确,面试评价人员一看就懂,不是纯理论层面提出的一些难以把握的抽象概念;适度指的是考核的指标数量要适当,既不能太多过于繁杂,要考虑到面试专家评价

时的工作量,但也不能太少导致覆盖面不足,要能够体现对研究生的综合素质和能力的要求。

3. 客观公正性和独立性原则

评价指标体系应该能够较为客观地反映出考生的真实水平,考核指标对于所有考生来说是公平公正的,不能因个别考生的特长设指标;设计的具体指标既要保证其全面性,能够考察考生综合素质的本质,又要让各指标之间保持相互独立,互不存在包含或交叉关系。

(二) 评价指标设计

按照以上指标设计原则,同时结合相关研究文献^[4-8]和对南京财经大学不同学院近30位研究生导师就招收学生时所看重学生的能力素质的访谈结果,制定了研究生面试的评价指标体系(见表1),面试时主要从专业素养和专业技能、科研潜力和创新能力、外语运用能力、逻辑思维和沟通能力四个维度进行考核,同时根据访谈情况给出了考核指标的建议权重占比(在面试操作过程中各专业可以根据各自的学科特点进行适当调整)。当然,除了这四个维度以外,研究生复试时还要考察考生的思想政治素质和道德品质、身体素质。思想政治素质和道德品质通常由人事档案审查或政审由学校研究生院统一完成,身体素质通常由医院体检认定,因此这两个维度暂不放在面试考察的评价指标体系当中。

表1 研究生面试评价指标体系

面试考核指标	考察重点	建议考核占比
专业素养和专业技能	大学阶段学习情况及成绩;考生对本学科理论知识和应用技能掌握程度。	25%
科研潜力和创新能力	利用所学理论发现、分析和解决问题的能力,对本学科发展动态的了解以及在本专业领域发展的潜力,好奇心、进取心和责任感等。	30%
外语运用能力	听力和口语表达能力,能够熟练运用英语进行沟通表达和学术研究。	25%
逻辑思维和表达沟通能力	正确合理地思考开放性问题并准确有条理地沟通表达等能力。	20%

专业素养和专业技能侧重考察考生在大学阶段的课程学习情况及成绩,根据考生提供的个人简历、课程成绩单等材料进行提问,了解考生对本学科理论知识和应用技能掌握程度。科研潜力的大小和创新能力强弱对于研究生的培养质量来说极为关键,是面试考核的重中之重,在面试过程中要考察考生对本学科发展动态的了解程度,判断其在本专业领域发展的潜力,通过提问了解考生利用所学理论发现、分析和解决问题的能力,也可以参考考生发表

与本专业领域相关的学术论文、获奖作品或者挑战杯、创新创业大赛、数学建模大赛、程序设计大赛等竞赛获奖情况以及参与的科研项目情况等方面进行判断;另外,在访谈过程中多位研究生导师提出,是否具备对问题的好奇心、事业进取心和责任感也直接影响学生在研究生阶段科研创新能力的发挥,在面试过程中也要加以考察。随着高校学术研究和人才培养的国际化程度不断提高,对研究生外语水平的要求也在不断提高,外语运用能力考察主要是在

面试过程中由专家用英语提几个问题考生现场用英语回答,考察面试对象的英语听力水平和口语沟通表达技能。逻辑思维和表达沟通能力未必要局限在专业领域的知识考察方面,可以采用开放性问题的考察考生运用科学合理的逻辑方法进行思考以及是否能够准确有条理地将其个人想法表达出来的能力。

三、基于公平公正视角的专家群体决策方法

根据研究生面试的测评要求,各学科点应成立复试面试小组,通过专家群体决策的方法给研究生进行评分,通常情况下所采取的做法是默认各专家的权重相等然后结合指标权重对评价结果进行集结。但正如前文所述,专家组在评价的过程中可能会有个别专家因为某些特殊原因导致对同一考生评价的结果与面试小组中其他人的评价结果差异性较大,如何保证面试小组专家评价的公平公正性是专家群体决策过程中非常重要的问题,因此本文设计了一种基于专家评价意见贴进度动态调整专家权重的方法,其思想是专家的实际决策权重将根据其个人意见与集体意见的贴进度进行动态调整。如果某个专家的评价意见与群体的评价意见偏差较大时,该专家的决策权重将被调低,相反,如果某个专家与集体意见的一致性程度越高,则该意见越能代表集体的看法,其意见的实际权重占比将在计算专家组最后的群体评价结果时被调高。

假设有 m 位专家组成研究生复试面试小组对 n 位考生进行面试,每位专家的评价初始权重 w_k 均等于 $\frac{1}{m}$,专家 k ($k=1, \dots, m$) 对于考生 i 的考核指标 j ($j=1, \dots, 4$) 评分为 u_{ij}^k ,考核指标的权重为 f_j ,这样专家 k 对 n 位考生进行评价最后形成的评价矩阵为 $D_k = (u_{ij}^k)_{n \times 4}$; m 位专家的评价意见集成后形成的群体评价矩阵为 $D = (u_{ij})_{n \times 4}$ [9],其中 $u_{ij} =$

$$\sum_{k=1}^m u_{ij}^k \tau w_k。$$

定义 1 [10] 设 D' 为专家评价与群体评价的偏离距离矩阵, $D' = (u'_{ij})_{n \times 4}$, 其中 $u'_{ij} = \sqrt{\sum_{j=1}^4 (u_{ij} - u_{ij}^k)^2 f_j^2}$, 令向量 $u^* = (u_1^*, u_2^*, \dots, u_n^*)$ 和 $u^0 = (u_1^0, u_2^0, \dots, u_n^0)$ 分别为评价偏离距离矩阵中的正、负理想点,其中 $u_i^* = \min_{0 \leq j \leq 4} u'_{ij}$, $u_i^0 = \max_{0 \leq j \leq 4} u'_{ij}$, 即正理想点表示偏离距离中的最小值,负

理想点表示偏离距离中的最大值。

定义 2 [11] 设专家 k 与专家群体判断的正理想点距离为 d_k^* , 负理想点距离为 d_k^0 , 则专家 k 与专家群体判断理想点的贴进度 $t_k = \frac{d_k^0}{d_k^0 + d_k^*}$, 其中

$$d_k^* = \sqrt{\sum_{j=1}^n (c_{ij} - u_j^*)^2}, d_k^0 = \sqrt{\sum_{j=1}^n (c_{ij} - u_j^0)^2}。$$

根据专家意见的贴进度,对其权重进行动态调整,在此基础上进行专家群体评价意见的集结,具体算法的步骤如下:

Step 1 专家按照研究生面试考核指标的 4 个维度对 n 位考生逐一进行面试评分,形成评价矩阵 D_k ;

Step 2 根据 D_k 和专家的初始权重 w_k 求解面试专家组的群体评价矩阵 D ;

Step 3 根据定义 1 中的公式求解专家评价与群体评价的偏离距离矩阵 D' , 找出群体判断的正负理想点;

Step 4 根据定义 2 中的公式计算各专家评价与群体评价理想点的贴进度;

Step 5 根据贴进度调整专家决策权重并进行归一化处理,可得专家调整后的权重 w'_k ;

Step 6 由调整后的专家权重 w'_k 重新对专家的评价矩阵为 D_k 进行集结,再与相应的指标权重 f_j 进行加权,得到调整后的面试排序结果。

四、案例分析

某大学管理科学与工程学院为了推进 2018 年该院管理科学与工程专业的研究生复试工作,成立了研究生面试专家组,专家组成员 5 人,每位成员的评价意见初始权重均为 0.2,专家组对进入复试的 8 位考生进行面试并按照表 1 构建的评价指标体系对面试考生进行评分,专家给出的评价矩阵如下:

$$D_1 = \begin{bmatrix} 75 & 82 & 70 & 85 \\ 85 & 75 & 80 & 80 \\ 95 & 95 & 95 & 95 \\ 71 & 65 & 75 & 70 \\ 85 & 87 & 90 & 85 \\ 80 & 85 & 80 & 83 \\ 71 & 77 & 75 & 80 \\ 92 & 90 & 80 & 93 \end{bmatrix},$$

$$D_2 = \begin{bmatrix} 80 & 80 & 70 & 80 \\ 80 & 70 & 75 & 75 \\ 80 & 85 & 80 & 85 \\ 73 & 68 & 70 & 70 \\ 84 & 85 & 85 & 80 \\ 85 & 82 & 85 & 80 \\ 75 & 85 & 77 & 85 \\ 90 & 90 & 85 & 95 \end{bmatrix},$$

$$D' = \begin{bmatrix} 3.305 & 0.79 & 1.61 & 1.28 \\ 2.72 & 1.52 & 2.3 & 0.61 \\ 2.07 & 2.36 & 0.52 & 2.36 \\ 1.61 & 0.49 & 0.97 & 0.97 \\ 1.21 & 3.04 & 2.13 & 1.54 \\ 1.34 & 1.06 & 0.68 & 1.31 \\ 2.1 & 3.09 & 1.31 & 1.24 \\ 2.11 & 2.13 & 3.41 & 2.65 \end{bmatrix}$$

$$D_3 = \begin{bmatrix} 85 & 80 & 70 & 80 \\ 80 & 75 & 85 & 85 \\ 80 & 80 & 85 & 80 \\ 75 & 65 & 70 & 70 \\ 80 & 80 & 85 & 75 \\ 81 & 83 & 80 & 80 \\ 70 & 72 & 70 & 75 \\ 85 & 80 & 80 & 85 \end{bmatrix},$$

因此可计算专家组群体评价的正理想点向量 u^* 和负理想点向量 u^0 分别为: $u^* = (0.79, 0.567, 0.52, 0.49, 0.541, 0.68, 1.088, 1.135)$, $u^0 = (3.305, 2.72, 6.26, 1.878, 3.04, 1.34, 3.09, 3.41)$ 。

$$D_4 = \begin{bmatrix} 78 & 80 & 70 & 80 \\ 88 & 80 & 85 & 85 \\ 85 & 85 & 85 & 85 \\ 77 & 70 & 80 & 75 \\ 88 & 90 & 93 & 89 \\ 84 & 85 & 84 & 85 \\ 75 & 80 & 80 & 83 \\ 95 & 93 & 85 & 95 \end{bmatrix},$$

根据定义 2 中的公式计算各专家评价与群体评价的正、负理想点距离,并在此基础上可得各专家评价与群体评价理想点的贴适度分别为: 0.4118, 0.6217, 0.5440, 0.6633, 0.6691。根据贴适度调整专家决策权重并进行归一化处理,可得专家调整后的权重分别为: 0.1416, 0.2137, 0.1869, 0.2279, 0.2299。由调整结果可见,专家 1 的意见权重被降低,主要原因是其对考生 3 的评价结果与其他 4 位专家的差异性过大,其他各位专家的意见权重也和调整前相比有所变化。最后根据调整后的专家权重对评价结果进行集结,表 2 给出了专家权重调整前和调整后的考生评分排序情况对比,由表 2 可见,考生 3 的综合评价排名由第一名变成了第三名,调整前考生 1 排序为最后一名,而调整后考生 4 的排名最末,如果面试实行差额录取,则录取的结果和调整前发生了变化。

$$D_5 = \begin{bmatrix} 79 & 81 & 70 & 80 \\ 83 & 77 & 80 & 80 \\ 84 & 80 & 85 & 80 \\ 75 & 74 & 80 & 75 \\ 87 & 90 & 90 & 87 \\ 84 & 87 & 82 & 88 \\ 75 & 82 & 75 & 85 \\ 90 & 88 & 82 & 91 \end{bmatrix}$$

表 2 专家权重调整前后考生的排序结果

根据上一节的算法步骤,可得面试专家组的群体评价矩阵 D 以及专家评价与群体评价的偏离距离矩阵 D' ,其中:

$$D = \begin{bmatrix} 79.4 & 78.2 & 73.2 & 78 \\ 83.2 & 75.4 & 81 & 81 \\ 85.4 & 85.8 & 86 & 86.8 \\ 75.6 & 77.4 & 79 & 77 \\ 84.8 & 86.4 & 88.6 & 83.2 \\ 82.8 & 84.4 & 82.2 & 83.2 \\ 73.2 & 79.2 & 75.4 & 81.6 \\ 87 & 85.6 & 82.4 & 89 \end{bmatrix},$$

考生序号	权重调整前 综合评分	权重调整 前排序	权重调整后 综合评分	权重调整后 后排序
1	77.21	8	77.51	7
2	79.87	5	79.92	5
3	85.95	1	85.12	3
4	77.27	6	77.47	8
5	85.91	2	86.10	1
6	83.21	4	83.40	4
7	77.23	7	77.56	6
8	85.83	3	86.10	2

五、结束语

研究生考试作为国家选拔人才最重要的考试之一,其公平公正性一直受到社会各界的强烈关注,面试作为研究生复试当中主观性比较大的一个考核环节,如何保证面试结果的公平公正性就显得尤为重要。研究生面试的过程是一个群体决策的过程,其评价指标体系和决策方法既要科学合理又要简单实用,本文提出了研究生面试的主要考核指标,并在我们经常使用的各位专家按照相同权重进行加权的算法基础上,提出了一种基于专家评价意见贴近度动态调整专家权重的改进方法,案例计算的结果表明,其可以一定程度上改变个别专家由于个人的某种原因给个别考生评分过高或者过低所造成的不公正结果,也使得专家在评分时更加慎重,更多地站在公平公正和全局的角度进行评价,进而形成面试评价工作的良性循环。考虑到后期应用的便捷性,面试之前可通过 MATLAB 将本算法进行编程,面试结束后将专家的评分输入初始矩阵,就可以直接得到最终的计算结果。

参考文献:

[1] 魏峻,姬红兵,纪树东.机会公平视角下的研究生复试工

作再认识[J].研究生教育研究,2017(3):31-34.

- [2] 肖文英,孙秋红,孙乐平,等.硕士研究生入学考试复试中科学性 with 公平性问题研究[J].学位与研究生教育,2015(2):56-60.
- [3] 司伟建,张朝柱,穆琳琳,等.影响硕士研究生复试“结构化面试”中面试官评分公平性的因素及解决策略[J].教育教学论坛,2018(1):1-2.
- [4] 孙晓敏,车宏生.关于研究生复试考核维度的思考——国外研究生胜任特征研究的启示[J].中国高教研究,2012(9):40-48.
- [5] 朱辉荣,王平义,白榕.研究生创新素质测评体系在复试中的构建[J].黑龙江高教研究,2012(6):141-143.
- [6] 郑艳,段亚敏.全日制专业学位硕士研究生招生复试评价体系研究[J].石家庄经济学院学报,2015,38(6):136-141.
- [7] 高静.硕士研究生复试指标体系的构建研究——以燕山大学为例[J].文教资料,2015(2):93-96.
- [8] 王传毅,程哲.研究生招生考试中“非认知能力”的测量:概念、实践与展望[J].研究生教育研究,2017(5):67-72.
- [9] 孙晓东,田澎.群决策中基于一致性强度的专家意见集结方法[J].系统工程与电子技术,2008,30(10):1895-1898.
- [10] 唐润,王海燕.基于直觉模糊-TOPSIS 群决策方法的白酒质量评价[J].系统工程,2012(8):97-100.
- [11] 王俊英,李德华.群决策专家权重自适应算法研究[J].计算机应用研究,2011,28(2):97-100.

Research on Interview-related Problems for Admitting Postgraduates from the Perspective of Fairness and Justice

TANG Run

(School of Management Science & Industrial Engineering, Nanjing University of Finance and Economics, Nanjing 210023)

Abstract: Admission interview is a very important chain for finding excellent candidates in the postgraduate enrollment. How to ensure fairness and justice of the interview is a critical issue worth studying. This paper first analyzes the main factors that affect the fairness and justice of the admission interviewers, and then based on the analysis, proposes an evaluation index system for the outcomes of the postgraduate interview, putting forward a method to dynamically adjust expert weight based on the close degree of the evaluation of the experts. Finally, this paper analyzes the correction capacity for evaluation by applying the algorithm in case studies and comparison, and verifies that the method can reduce, in a certain degree, the probable variation in rigidity of some experts in evaluating a candidate during the interview.

Keywords: fairness and justice; postgraduate enrollment; weight; interview