

福建省青少年体操锦标赛 E 组评分客观性分析 ——以 2021 年男子体操单项决赛为例

李明,王丰彩

(集美大学诚毅学院,福建 厦门 361021)

摘要:研究通过对 2021 年福建省青少年体操锦标赛男子单项决赛 E 组裁判评分数据进行客观性分析,探讨了福建省体操裁判评分的准确度和客观性,并提出相关建议,旨在提高福建裁判团队的执裁水平。结果表明:三个组别 6 个单项决赛项目的 E 组裁判员评分的客观性系数除了丁组的吊环和自由体操两个项目的一致性系数稍低外,其他项目的评分一致性系数都比较高,裁判员评分的随机误差情况整体上控制较好;三个组别 6 个项目的评分差异不显著,裁判员评分的标准比较一致;个别裁判员极差值和无效分的出现机率稍大,但对运动员最后得分的可信度影响不大。建议:提高福建体操裁判执裁评分标准的准确度和一致性,合理控制和处理评分误差,降低无效分的出现机率,确保男子体操裁判评分的客观性,完善体操裁判评价机制。

关键词:青少年锦标赛;体操比赛;裁判员评分;客观性

中图分类号:G832

文献标识码:A

文章编号:1007-7413(2022)05-0052-05

Objective Analysis of Referees' Scoring in Group E of Fujian Provincial Youth Gymnastic Championship

——Case Study on the Men's Individual Gymnastics Final in 2021

LI Ming, WANG Feng-cai

(Chengyi College, Jimei University, Xiamen 361021, China)

Abstract: In ways of literature review, field observation and mathematical statistics, the objective analysis of the scoring data of Group E judges of the men's single event final of the 2021 Fujian Provincial Youth Gymnastics Championship, this paper investigated and analyzed the accuracy and objectivity of the scoring of gymnastics judges in Fujian Province, then put forward relevant suggestions to improve the adjudication level of Fujian referees. The preliminary findings are as follows: the objectivity coefficient of the judges' scores in group E of the six individual final events of the three groups is relatively high, except for the consistency coefficient of the rings and floor exercise of group D, which is slightly lower. The random error of the judges' scores is well controlled on the whole. There was no significant difference in the scores of the 6 items in the three groups, and the judges' scoring standards were consistent. The probability of range and invalid score of individual judges is slightly larger, but it has little influence on the credibility of the final score of athletes. Therefore, to ensure the objectivity of the scoring of men's gymnastics judges and improve the evaluation mechanism of gymnastics judges, we should improve the accuracy and consistency of the scoring standards of Fujian gymnastics judges, reasonably control and deal with the scoring errors, reduce the probability of invalid scores.

Key words: Youth Championship; gymnastics competition; judge's score; objectivity

体操是我国体育的传统优势项目,不仅具有观赏性,而且也广受大众的喜爱。福建省尤为重视该项目的发展,并将体操项目列入后备人才培养模式创新试点,在运动员、教练员、裁判员培养机制上给予大力支

持。裁判员是体操比赛中的执法官,裁判员评分的准确性和客观性对比赛的结果有着关键性的影响^[1],决定了比赛的公平、公正性。2021 年福建省青少年体操锦标赛在泉州海峡体育中心体育馆举行,共有来自福

收稿日期:2021-12-20

第一作者简介:李明(1978—),男,山东费县人,讲师,硕士。研究方向:体育教育训练学。

建省内各设区市、平潭综合试验区、省重点传统校的9个代表队,总计169名运动员参赛。本届比赛采用《2017—2020年国际竞技体操评分规则》和《中国青少年体操训练教学大纲》中附件1的“中国少年儿童体操比赛规定动作及评分标准”。在男子竞技体操评分规则中,D组裁判打难度分,E组裁判打完成分,两组裁判给分均是独立的,去掉最高分和最低分,最后的平均分即为参赛队员的最后得分。因同年年龄段的参赛队员难度分差异性不是很大,因此,E组裁判评分的客观性对比赛结果起着决定的作用。本研究以2021年福建省青少年体操锦标赛男子单项决赛成绩中E组裁判的评分作为研究对象,对裁判员评分的客观性进行分析,以此为依据,了解福建省体操裁判评分的准确性、客观性和执裁评分的经验,对裁判员评分机制的公平性研究和裁判员的遴选、培养工作具有一定的现实意义。研究分析并提出相应对策,以期提高福建体操裁判执裁水平,保证福建各类体操赛事的成绩准确性和竞赛公正性,对运动员和教练员负责,进一步促进福建体操项目竞技水平和赛事推广的健康发展。

1 研究对象与方法

1.1 研究对象

以2021年福建省青少年体操锦标赛乙组、丙组、丁组男子鞍马、单杠、吊环、跳马、双杠、自由操的6个单项决赛前八名运动员的完成情况E分为研究对象(因甲组参赛运动员较少,不作为本次研究对象)。

1.2 研究方法

1.2.1 文献资料法

通过文献资料法查阅中国知网和2016年国际体操联合会男子体操评分规则,结合2021年福建省青少年锦标赛成绩册对相关研究文献进行内容梳理,为论文研究提供有力的理论支撑和参考。

1.2.2 实地观察法

对2021年福建省青少年体操锦标赛中的男子单项决赛6个项目中的运动员完成情况和赛事运行情况进行实地观察,便于结合赛场上运动员的临场发挥及动作表现研究裁判员的评分情况。

1.2.3 数理统计法

运用统计软件SPSS23.0中的相关性分析、方差分析和极差分析,结合Microsoft Office Excel2010对2021年福建省青少年体操锦标赛男子单项决赛E组裁判评分数据进行统计分析,对裁判员执裁客观性进

行综合评价。

2 结果与分析

2.1 客观性检验

应用SPSS23.0统计软件,对2021年福建省青少年体操锦标赛男子体操单项决赛中E组裁判员对前八名运动员动作完成质量的评分进行肯德尔和谐系数(Kendall's W)处理和计算,即检验其一致性系数,客观反映男子体操单项决赛中3个组别6个项目的E组裁判员评分的一致性情况。

从表1、表2、表3中男子乙组、丙组、丁组单项决赛6个项目检验统计结果可以看出,渐进显著性 P 值均小于显著水平0.05,Kendall's W检验结果中除丁组的吊环一致性系数为0.577和自由体操一致性系数为0.503之外,其他组别的各项目评分一致性系数都在0.7以上。从统计学角度来看,男子乙组、丙组、丁组6个单项决赛中运动员个体的得分存在显著差异,说明三个组别的E组裁判员的整体评分标准一致性较高。输出的分析结果说明:比赛中运动员动作差异区分度较高,裁判员的评分虽然有一定的误差,但是相比运动员之间显著的动作技术差异,评分的误差对运动员最终得分的影响不大。从个案样本一致性结果来看,男子乙组双杠项目的E组裁判评分一致性系数最高为0.973,说明乙组双杠裁判的评分客观性最高;而在丙组和丁组双杠项目中,裁判评分的一致性系数分别为0.940和0.896,也相对较高,反映了在双杠项目中裁判评分标准一致,现场评分较为精准。男子丁组的吊环和自由体操两个项目的E组裁判评分一致性系数相对较低,相对其他决赛项目来看,丁组吊环和自由体操两个项目的一致性数据反应了该项目中E组裁判的业务水平存在差异,评分的标准和尺度不一致。另外,丁组的自由体操项目评分的渐进显著性 P 值为0.014,大于显著性水平0.01而小于显著性水平0.05,虽然渐进显著性水平具有显著差异,但结合一致性系数检验结果也反应出该项目的E组裁判对该项目评分尺度的把控不准确,整体业务水平有待提高。从现场观察及成绩册上的评分数据来看,丁组决赛的吊环项目中难度分均相同,而丁组自由体操项目中除了一名队员的难度分为4.4分,其余7名运动员的难度分均为5分,所以在这2项中,完成分即E分的高低基本决定了运动员最终成绩的高低,E分的不客观会直接对比赛结果造成不良影响。另外,项目的动作特征也是影响评

分一致性的重要因素,单杠、双杠项目动作有动有静,过程较为舒缓,有利于裁判的仔细观察和思考斟酌,因此以上项目裁判员评分较准确,客观一致性也高;而跳马和自由体操等项目,其动作较为连贯,甚至是瞬间一气呵成,留给裁判员观察思考的时间很少,这对裁判员的执裁水平和经验是很大的考验。

表 1 男子乙组单项决赛 E 组裁判员评分的一致性系数检验分析

项目	Kendall's Wa	Chi-Square	P
鞍马	0.765	22.949	0.001
单杠	0.918	22.943	0.000
吊环	0.863	30.188	0.000
双杠	0.973	24.314	0.000
跳马	0.799	27.966	0.000
自由体操	0.845	29.588	0.000

注: $P < 0.01$ 差异非常显著; $P < 0.05$ 差异显著; $P > 0.05$ 差异不显著。

表 2 男子丙组单项决赛 E 组裁判员评分的一致性系数检验分析

项目	Kendall's Wa	Chi-Square	P
鞍马	0.876	17.510	0.002
单杠	0.919	27.561	0.000
吊环	0.934	28.007	0.000
双杠	0.940	28.200	0.000
跳马	0.706	21.174	0.002
自由体操	0.946	28.366	0.000

注: $P < 0.01$ 差异非常显著; $P < 0.05$ 差异显著; $P > 0.05$ 差异不显著。

表 3 男子丁组单项决赛 E 组裁判员评分的一致性系数检验分析

项目	Kendall's Wa	Chi-Square	P
鞍马	0.895	31.333	0.000
单杠	0.900	31.508	0.000
吊环	0.577	20.203	0.005
双杠	0.896	31.348	0.000
跳马	0.808	28.271	0.000
自由体操	0.503	17.601	0.014

注: $P < 0.01$ 差异非常显著; $P < 0.05$ 差异显著; $P > 0.05$ 差异不显著。

2.2 方差检验

为了检验裁判员评分标准对运动员 E 分这个独

立因变量均值是否存在显著差异,采用 SPSS 统计软件中的单因素方差分析对 2021 年福建省青少年体操锦标赛男子体操单项决赛中前 8 名运动员动作完成质量情况进行了检验,通过方差分析输出情况得出男子体操单项决赛中 6 个项目的方差齐性检验统计量。

从表 4 的方差同质性检验分析输出可以看出:丙组的跳马乙组、丙组、丁组 3 个组别中的 6 个项目的 E 组裁判员评分方差同质性检验显著性均大于显著水平 0.05,这说明 3 个组别中 6 个项目的 E 组裁判员评分数据的总体方差不存在显著差异。再结合 Levene 统计量来看,可以认为 3 个组别中 6 个项目的总体方差相等,即满足方差齐性这个前提条件,可以进行计算分组均值相等的 Levene 统计量,进一步检验组方差是否相等和各单项 E 组裁判员评分的一致性方差分析情况。

表 4 男子单项决赛 6 个项目的方差同质性检验分析

项目	裁判评分数据					
	乙组		丙组		丁组	
	Levene 统计	显著性	Levene 统计	显著性	Levene 统计	显著性
鞍马	0.206	0.933	0.566	0.690	0.464	0.761
单杠	0.481	0.750	0.269	0.896	1.563	0.206
吊环	2.597	0.053	0.723	0.583	2.030	0.111
双杠	0.425	0.789	1.059	0.394	1.051	0.395
跳马	0.602	0.664	2.819	0.083	1.510	0.221
自由体操	0.572	0.684	0.344	0.846	1.884	0.135

从单因素方差分析输出情况 6 个项目单项决赛可以看出(表 5、表 6、表 7),如果仅考虑裁判员单个因素的影响,3 个组别中 6 个项目的 E 组裁判员的评分标准没有显著差异($P > 0.05$),即整体裁判员的评分标准和尺度较为一致,可以认为 3 个组别中 6 个项目的 E 组裁判员评分标准差异较小,没有对运动员决赛成绩产生显著影响,能客观反映参赛运动员的真实水平。同时也体现出:虽然裁判员对所有运动员的总体评分尺度较为一致,没有倾向性的偏高或偏低,但对于水平相近的运动员来讲,裁判员评分的区分度不高,仍有待进一步提高辨识度。^[1]造成这种结果的原因可能有两方面:一方面从客观角度看,运动员运动水平较为接近,裁判员在评分时难以评判出高低;另一方面从主观角度来看,裁判员在评分过程中相对保守,为让自己的评分成为有效分,评分的波动范围较小,从而降低了区分度。^[2-8]

表 5 男子乙组单项决赛各单项 E 组裁判员评分一致性
方差分析

项目		SS	df	MS	F	P
鞍马	组间	1.670	4	0.417	0.845	0.508
	组内	14.826	30	0.494		
	总和	16.495	34			
单杠	组间	0.201	4	0.050	0.054	0.994
	组内	23.368	25	0.935		
	总和	23.570	29			
吊环	组间	0.292	4	0.073	0.393	0.812
	组内	6.508	35	0.186		
	总和	6.800	39			
双杠	组间	0.440	4	0.110	0.216	0.927
	组内	12.707	25	0.508		
	总和	13.147	29			
跳马	组间	0.168	4	0.042	0.780	0.546
	组内	1.891	35	0.054		
	总和	2.060	39			
自由体操	组间	0.655	4	0.164	0.381	0.821
	组内	15.040	35	0.430		
	总和	15.695	39			

注: $P<0.01$ 差异非常显著; $P<0.05$ 差异显著; $P>0.05$ 差异不显著。

表 6 男子丙组单项决赛各单项 E 组裁判员评分一致性
方差分析

项目		SS	df	MS	F	P
鞍马	组间	0.162	4	0.040	0.132	0.969
	组内	6.112	20	0.306		
	总和	6.274	24			
单杠	组间	0.839	4	0.210	0.532	0.713
	组内	11.820	30	0.394		
	总和	12.659	34			
吊环	组间	0.567	4	0.142	0.331	0.855
	组内	12.860	30	0.429		
	总和	13.427	34			
双杠	组间	1.987	4	0.497	0.304	0.873
	组内	49.040	30	1.635		
	总和	51.027	34			
跳马	组间	0.090	4	0.022	0.067	0.991
	组内	10.117	30	0.337		
	总和	10.207	34			
自由体操	组间	1.581	4	0.395	0.508	0.731
	组内	23.366	30	0.779		
	总和	24.947	34			

注: $P<0.01$ 差异非常显著; $P<0.05$ 差异显著; $P>0.05$ 差异不显著。

表 7 男子丁组单项决赛各单项 E 组裁判员评分一致性
方差分析

项目		SS	df	MS	F	P
鞍马	组间	1.699	4	0.425	0.408	0.802
	组内	36.451	35	1.041		
	总和	38.150	39			
单杠	组间	0.811	4	0.203	0.528	0.716
	组内	13.445	35	0.384		
	总和	14.256	39			
吊环	组间	0.641	4	0.160	2.021	0.113
	组内	2.777	35	0.079		
	总和	3.419	39			
双杠	组间	1.751	4	0.438	0.396	0.810
	组内	38.706	35	1.106		
	总和	40.458	39			
跳马	组间	0.529	4	0.132	1.208	0.325
	组内	3.827	35	0.109		
	总和	4.356	39			
自由体操	组间	0.289	4	0.072	0.398	0.808
	组内	6.349	35	0.181		
	总和	6.638	39			

注: $P<0.01$ 差异非常显著; $P<0.05$ 差异显著; $P>0.05$ 差异不显著。

2.3 极差分析

为寻找影响 E 组裁判员评分一致性的主要因素,采用 SPSS23.0 统计软件中的描述统计分析,并结合 Microsoft Office Excel2003 软件,运用极差分析对 2021 年福建省青少年体操锦标赛男子单项决赛 E 组裁判评分出现极差的频率进行分析统计,结果如表 8、表 9、表 10 和图 1 所示。

表 8 男子乙组单项决赛各单项 E 组裁判员扣分出现
极差的频率(次)

极差/分	鞍马	单杠	吊环	双杠	跳马	自由体操	合计
0.1	0	0	0	1	2	1	4
0.2	0	0	1	1	3	0	5
0.3	0	0	0	1	1	1	3
0.4	0	1	3	0	1	0	5
0.5	0	0	0	2	1	0	3
0.6	3	0	0	0	0	2	5
0.7	1	2	2	0	0	0	5
0.8	1	1	1	0	0	2	5
0.9	0	2	0	0	0	1	3
1.0	1	0	1	0	0	1	3
1.1	1	0	0	0	0	0	1
1.6	0	0	0	1	0	0	1

表 9 男子丙组单项决赛各单项 E 组裁判员扣分出现极差的频率 (次)

极差/分	鞍马	单杠	吊环	双杠	跳马	自由体操	合计
0.2	0	0	0	1	2	0	3
0.3	0	2	1	0	0	0	3
0.4	0	0	2	0	0	0	2
0.5	2	2	0	1	1	0	6
0.6	1	1	1	1	3	0	7
0.7	2	0	2	1	1	2	8
0.8	0	1	0	0	0	4	5
0.9	0	1	1	0	0	1	3
1.2	0	0	0	2	0	0	2
1.9	0	0	0	1	0	0	1

表 10 男子丁组单项决赛各单项 E 组裁判员扣分出现极差的频率 (次)

极差/分	鞍马	单杠	吊环	双杠	跳马	自由体操	合计
0.1	0	0	0	0	1	0	1
0.2	0	1	0	0	1	0	2
0.3	0	1	2	0	2	1	6
0.4	0	0	0	1	0	0	1
0.5	1	1	1	0	2	1	6
0.6	0	1	0	1	0	1	3
0.7	0	0	4	1	1	1	7
0.8	1	1	1	1	1	1	6
0.9	3	1	0	0	0	1	5
1.0	1	0	0	1	0	1	3
1.1	0	2	0	0	0	0	2
1.2	0	0	0	0	0	1	1
1.3	1	0	0	0	0	0	1
1.4	0	0	0	1	0	0	1
1.5	1	0	0	1	0	0	2
1.9	0	0	0	1	0	0	1

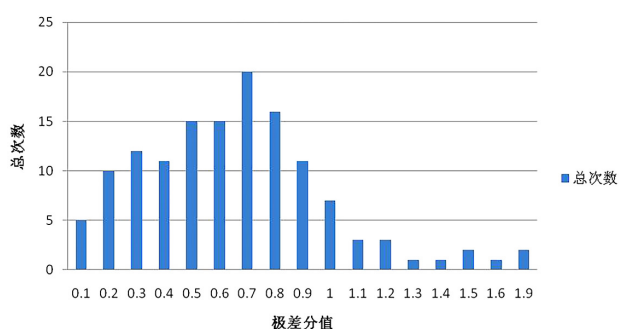


图 1 E 组裁判评分极差分值与出现总次数对应图

通过表 8、表 9、表 10 对男子体操单项决赛 E 分的分析和统计可以看出:E 组裁判对运动员的评分极差在 0.1~1.9 之间波动。结合图 1 的 E 组裁判评分极差分值与出现总次数对应图可以看出:汇总 3 个组别

的评分极差值和出现次数,最小评分极差为 0.1,在 3 个组别比赛中共出现了 5 次;最大评分极差为 1.9,分别在丙组和丁组双杠决赛中各出现了 1 次;出现最多的评分极差为 0.7,在比赛中出现了 20 次;所有极差值在不同位置出现的次数大致成偏正态分布,符合统计学的要求。极差值的出现机率和位置说明了裁判员个体间评分存在差异,有一定的随机误差,而且在部分项目中,裁判员的评分差异较大。^[2]例如在丙组和丁组双杠决赛项目的评分中,裁判的评分极差值为 1.9 分,这说明裁判员对个别运动员的评分争议较大,评分数据的客观性受到质疑。另外,3 个组别中的最高极差都出现在双杠项目中,这说明该项目中个别裁判员评分出现较大的随机误差,评分的精准度和执裁水平有待于进一步提高。虽然裁判员的评判结果出现明显的个体间差异,但是依据体操竞赛规则去掉最高分和最低分,仅出现一次差异显著的评分并不会影响整体裁判员的最终评分结果。

3 结论

(1)青少年体操锦标赛男子体操单项决赛中,大部分单项的 E 组裁判员的整体评分客观一致性较高,对评分尺度的把握相对较为一致。但个别单项如吊环、跳马、自由体操等,评分存在一定的差异性,分析其原因主要是裁判团队的执裁水平不一和项目的动作特征差异。

(2)各个项目的裁判员评分的客观一致性系数虽然较高,但通过评析裁判员评分客观性的一系列指标数据后发现:个别裁判员评分的随机误差稍大,裁判员队伍水平参差不齐,部分裁判执裁经验不足,执裁水平有待提高。

(3)极差值和无效分的出现机率和出现位置,也反映了个别裁判员评分标准出现偏差,虽然这会使评分数据的客观性受到轻微质疑,但并不影响最终比赛成绩的可信度和评分结果。

(4)裁判员对所有运动员的总体评分标准和尺度较为一致,但对于水平相近的运动员来讲,裁判员评分的区分度不高,仍有待进一步提高辨识度。

4 建议

(1)优化裁判团队培养模式,实施裁判员继续教育
(下转第 87 页)

- 之道[J]. 高等教育研究,2020,41(6):26-33.
- [19]王家宏,韩春利. 中国特色体育学科体系的构建与完善[J]. 上海体育学院学报,2021,45(8):1-10.
- [20]周萍. 大学女性青年教师的角色冲突分析——社会性别视角下对A大学的质性研究[J]. 高校教育管理,2016,10(3):99-105.
- [21]王焕,闫静. 女性体育科研人员的科研产出现状与影响因素——以2000—2015年《体育科学》载文作者为例[J]. 上海体育学院学报,2017,41(5):42-48.
- [22]刘新民,俞会新. 高校青年教师科研压力对科研绩效的影响研究——基于认知评价视角[J]. 北京社会科学,2018(10):76-88.
- [23]李强. 制约体育院系青年教师参与科研活动的瓶颈主因子分析——基于马斯洛需要层次理论的研究[J]. 南京体育学院学报(社会科学版),2017,31(2):122-128.
- [24]科技部. 科技部印发《关于破除科技评价中“唯论文”不良导向的若干措施(试行)》的通知[EB/OL]. (2020-02-23) [2022-03-25]. http://www.most.gov.cn/xxgk/xinxifenlei/fdzdgknr/fgzc/gfxwj/gfxwj2020/202002/t20200223_151781.html.
- [25]科技部. 关于政协十三届全国委员会第三次会议第2415号(科学技术类130号)提案答复的函[EB/OL]. (2020-09-15) [2022-03-25]. http://www.most.gov.cn/xxgk/xinxifenlei/fdzdgknr/jyta/202101/t20210129_172660.html.
- [26]中共中央,国务院. 中共中央 国务院印发《深化新时代教育评价改革总体方案》[EB/OL]. (2020-10-13) [2022-03-25]. http://www.gov.cn/zhengce/2020-10/13/content_5551032.htm.
- [27]苏宴锋,司虎克. 体育科研核心作者群体特点及演变的计量分析[J]. 西南师范大学学报(自然科学版),2015,40(2):122-127.

[责任编辑 江国平]

(上接第56页)

育机制,对现有裁判团队定期组织业务培训,提高现有裁判员执裁水平。可利用理论讲解、案例探讨、对比赛视频模拟评分、实际操作等手段,强化业务练习,提高裁判员评分的准确度和辨识度,丰富执裁经验,在比赛中有效、合理地控制和处理评分误差,降低无效分的出现机率,提高比赛评判的准确性。

(2)强化裁判员的思想政治教育,在比赛中严格落实回避制度,同时加强监督监控机制,完善奖惩措施,出现重大过错时应追责,排除执裁过程中的主观因素干扰,使得裁判的评分更加公正、客观、准确,提高比赛评判的公正性。

参考文献

- [1]徐瑞芳. 北京奥运会男子体操单项决赛E组评分研究[J]. 体育科学研究,2009,13(3):50-51.
- [2]崔园园,郑幸红. 第29届奥运会女子体操单项完成分的客观性[J]. 体育科学研究,2010,14(2):71-73.
- [3]李萍. 大学生健美操比赛裁判员评分客观性分析:以重庆市第5届大学生健美操比赛为例[J]. 成都体育学院学报,2009,35(4):59-61.
- [4]王丰彩,杨雪红. 大学生运动会动感啦啦队比赛裁判员评分的客观性研究[J]. 西昌学院学报,2009,23(3):105-108.
- [5]张华,娄必刚. 2011年全国男子跳水锦标赛裁判员评分的客观性分析[J]. 西南师范大学学报,2012,37(12):152-155.
- [6]张晓莹,王宏,赵轩立. 竞技健美操裁判员执裁评估系统的研制与应用[J]. 北京体育大学学报,2015,38(10):120-125.
- [7]张晓莹,马全军,许寿生,赵轩立. 竞技健美操裁判员执裁水平的评价研究[J]. 中国体育科技,2015,51(6):62-68.
- [8]刘浩,史雨梅,余晓美. 基于SPSS单因素方差分析在专业认同研究中的应用[J]. 经济研究导刊,2020,26(5):71-73.

[责任编辑 江国平]